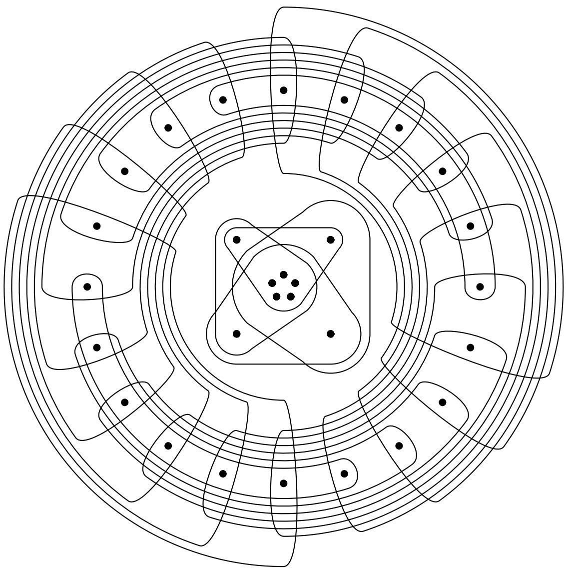
Luca Gugelmann

Ramsey properties of random graphs and hypergraphs



Diss. ETH No. 21394 2013

DISS. ETH NO. 21394

Ramsey properties of random graphs and hypergraphs

A dissertation submitted to
ETH ZURICH

for the degree of
DOCTOR OF SCIENCES

presented by
LUCA GUGELMANN
MSc ETH Math.
born 10.04.1984
citizen of Brittnau AG

accepted on the recommendation of
Prof. Dr. Angelika Steger, examiner
Prof. Dr. Mathias Schacht, co-examiner

2013

To my family

Contents

Abstract

The topic of this thesis are Ramsey-type problems in random graphs and hypergraphs. Ramsey theory has its origins in a famous 1930 paper by Frank P. Ramsey. Loosely speaking its object of study are large structures and conditions under which coloring these structures ensures that certain monochromatic substructures must always appear. A simple example is the question for the smallest complete graph such that regardless of how we color its edges with 2 colors, a monochromatic triangle always appears. The answer to this question is the complete graph K_6 on 6 vertices. In fact, even removing just a single edge from a K_6 makes it 2 colorable without a monochromatic triangle.

Random graphs were first introduced by Erdős and Rényi, and independently Gilbert, in 1959. Erdős and Rényi first considered the so-called $G(n, m)$ model, in which we select uniformly at random among all graphs with n vertices and m edges. Gilbert considered the $G(n, p)$ random graph on n vertices in which every possible edge is present with probability p independently.

Bollobás and Thomason proved in 1987 that all "nice" graph properties have a so-called *threshold*. A threshold is a function $p_0 = p_0(n)$ such that for $p \gg p_0$ it holds that $G(n, p)$ does have the property in question with probability tending to 1 for $n \to \infty$, while for $p \ll p_0$ the same

probability tends to 0. These two statements are usually referred to as the 1-statement and the 0-statement respectively. Much of the research in random graph theory has been devoted to finding exact values for the thresholds of various graph properties.

The study of Ramsey properties in random graphs was initiated by Łuczak, Ruciński and Voigt in 1992. For some fixed graph F and $r \geq 2$ they address the question of the threshold for the property that every r-edge-coloring of the random graph $G(n, p)$ contains a monochromatic copy of F. The problem was fully solved by Rödl and Ruciński which proved a 0-statement in 1993 and a matching 1-statement in 1995.

In a sense the results above can be seen as answering the question of how many random edges we need to remove from the complete graph on n vertices such that there exists an r-coloring without a monochromatic copy of a fixed graph F. In this setting we do not care which specific copy of F is monochromatic. A different kind of randomization was recently suggested by Allen, Böttcher, Hladký and Piguet. Instead of removing edges, they suggest to reduce the set of dangerous copies of F. The question is then for the number of copies of F which have to be randomly marked as not dangerous such that an r-coloring of the edges of K_n without a monochromatic copy of F which is still marked as dangerous becomes possible. In this thesis we address the question for the threshold of this problem, and also combine both kinds of randomization.

A k-uniform hypergraph is a generalization of the concept of graph in which every (hyper-)edge contains not 2 but k many vertices for some integer $k \geq 2$. A random k-uniform hypergraph on n vertices is such that every one of the $\binom{n}{k}$ possible hyperedges is present independently with probability p. For $k = 2$ this corresponds to the $G(n, p)$ model.

With intuitive arguments Rödl and Ruciński conjectured in 1998 that a natural extension of their results for the 1-statement in the graph case should also hold for random k-uniform hypergraphs. They proved this conjecture for 2 colors and the complete 3-uniform hypergraph on 4 vertices. Together with Schacht in 2007 they expanded this to all k-partite, k-uniform hypergraphs. In 2010 Friedgut, Rödl and Schacht, settled the general conjecture. Conlon and Gowers also independently obtained similar results.

It is widely believed that this 1-statement is tight for most hypergraphs F. In this thesis we prove a matching 0-statement showing that this is indeed true for many hypergraphs. However we also show that there are many more exceptions to this than in the graph case. In particular we show an example for which the 1-statement is not tight and threshold is given by the so-called *asymmetric* Ramsey game, in which a different hypergraph has to be avoided in each color.

We also turn our attention to so-called *online* problems. In every setting discussed so far we are allowed to see the entire random graph or hypergraph before committing to a coloring. In online games this is no longer the case and edges or vertices are revealed little by little in successive rounds, and we have to immediately commit to a coloring after each such round.

We study the online *balanced Ramsey game* in which a player has r colors and where in each step r random edges of an initially empty graph on n vertices are presented. The player has to immediately assign a different color to each edge and her goal is to avoid creating a monochromatic copy of some fixed graph F for as long as possible. We contrast this game to the similar *Achlioptas game* in which the player is again presented with r edges but is allowed to discard $r-1$ of them. Her goal is again to avoid creating a copy of F. This corresponds to a balanced Ramsey game in which only one color is forbidden. Similarly to the non-online (i.e. *offline*) cases, the typical duration of such a game exhibits a threshold-type behavior.

Krivelevich, Spöhel and Steger asked the question of whether these two games have the same threshold or not, as all results known so far show that this is the case for all non-trees. The intuition behind this question being that the balanced game may in fact not be more "difficult" than the Achlioptas game, as the player is always fighting some worst-case color and is not bothered by the remaining ones. Here we answer this question negatively for the edge-coloring version of these two games. We also consider the vertex-coloring variant and we show that in contrast to the edge case these two games indeed do have the same threshold.

Zusammenfassung

Das Thema dieser Arbeit sind Ramsey-Probleme in Zufallsgraphen und Zufallshypergraphen. Ramsey-Theorie hat ihren Ursprung in einer 1930 verfassten Arbeit von Frank P. Ramsey. Sie befasst sich im weitesten Sinne mit der Erforschung von Bedingungen die beim Färben einer grossen Struktur das Erscheinen von einfarbigen Substrukturen garantieren. Ein einfaches Beispiel ist die Frage nach dem kleinsten vollständigen Graphen, dessen Kanten man nicht mit zwei Farben so färben kann, dass kein einfarbiges Dreieck entsteht. Die Antwort zu dieser Frage ist der vollständige Graph K_6 mit 6 Knoten. Es gilt sogar, dass das Entfernen einer einzigen Kante genügt um eine 2-Färbung zu ermöglichen, welche kein einfarbiges Dreieck enthält.

Zufallsgraphen sind erstmals 1959 in den Arbeiten von Erdős und Rényi, und unabhängig davon die von Gilbert, erwähnt worden. Erdős und Rényi haben das sogenannte $G(n, m)$ Modell eingeführt. In diesem Modell wählt man uniform ein zufälliges Element der Menge aller Graphen mit n Kanten und m Knoten. Gilbert hingegen hat sich mit dem $G(n, p)$ Modell auseinandergesetzt, in dem man in einem Graphen mit n Knoten jede mögliche Kante unabhängig mit Wahrscheinlichkeit p auswählt.

Bollobás und Thomason haben 1987 bewiesen, dass "schöne" Eigenschaften von Graphen einen *Schwellenwert* haben. Ein Schwellenwert ist eine

Funktion $p_0 = p_0(n)$, so dass für $p \gg p_0$ gilt, dass $G(n,p)$ eine gewisse Eigenschaft mit Wahrscheinlichkeit $1 - o(1)$ besitzt, wogegen für $p \ll p_0$ diese Wahrscheinlichkeit gegen 0 strebt. Um einen solchen Schwellenwert genau zu bestimmen, beweist man üblicherweise eine untere und eine obere Schranke dafür. Ein grosser Teil der Forschung in diesem Gebiet beschäftigt sich dem genauen Bestimmen der Schwellenwerte für interessante Grapheneigenschaften.

Łuczak, Ruciński und Voigt haben 1992 als erste Ramsey-Probleme in Zufallsgraphen betrachtet. Für einen fixen Graphen F und $r \geq 2$ fragen sie nach dem Schwellenwert für die Eigenschaft, dass jede Färbung der Kanten von $G(n,p)$ mit r Farben eine einfarbige Kopie von F enthält. Diese Frage wurde von Rödl und Ruciński vollständig beantwortet, 1993 haben sie eine untere Schranke bewiesen und 1995 eine dazu passende obere Schranke.

In einem gewissen Sinne beantworten diese Resultate die Frage nach der Anzahl zufälliger Kanten, welche man aus einem vollständigen Graphen mit n Knoten entfernen muss, so dass eine r-Färbung möglich wird, welche keine einfarbige Kopie von F enthält. In diesem Zusammenhang ist es egal welche Kopie von F genau einfarbig ist, alle sind gleich gefährlich. Eine andere Art der Randomisierung wurde vor kurzem von Allen, Böttcher, Hladký und Piguet vorgeschlagen. Anstatt Kanten zu entfernen, könnte man auch die Menge der "gefährlichen" Kopien von F kleiner machen. Die Frage ist also nach der Anzahl Kopien von F welche man (zufällig) als ungefährlich markieren muss, um eine r-Färbung von K_n zu ermöglichen, welche keine immer noch gefährliche und einfarbige Kopie von F enthält. In dieser Arbeit bestimmen wir den genauen Schwellenwert für die Anzahl gefährlicher Kopien von F. Wir kombinieren auch die beiden obigen Arten der Randomisierung zu einer einzigen Aussage.

Ein k-uniformer Hypergraph ist eine Verallgemeinerung des Konzeptes eines Graphen. Jede Kante besteht nicht mehr aus 2, sondern aus k vielen Knoten, wobei $k \geq 2$ eine ganze Zahl ist. Ein k-uniformer Zufallshypergraph auf n Knoten enthält jede der $\binom{n}{k}$ möglichen Kanten unabhängig mit Wahrscheinlichkeit p. Für $k = 2$ entspricht dies dem $G(n,p)$ Modell.

Mit intuitiven Argumenten haben Rödl und Ruciński 1998 die Vermutung aufgestellt, dass ihre Aussagen über die obere Schranke für den Schwellenwert im Graphenfall auch in ähnlicher Form für Hypergraphen gelten müssten. Sie haben dies für den Fall von 2 Farben und dem vollständigen 3-uniformen Hypergraphen auf 4 Knoten bewiesen. Zusammen mit Schacht haben sie ihre Vermutung später auch für k-partite, k-uniforme Hypergraphen bewiesen. Ein Beweis dieser Aussage für beliebige Hypergraphen wurde 2010 von Friedgut, Rödl und Schacht gefunden. Conlon und Gowers haben unabhängig ähnliche Resultate erzielt.

Es wird allgemein angenommen, dass das obige Resultat für eine grosse Klasse von Hypergraphen bestmöglich ist. In dieser Arbeit beweisen wir dies mit einer passenden unteren Schranke. Wir beweisen aber auch, dass im Gegensatz zum Graphenfall für viel mehr Hypergraphen die obere Schranke von Rödl und Ruciński nicht bestmöglich ist. Insbesondere zeigen wir ein Beispiel von einem Schwellenwert der strikte unter deren oberen Schranke liegt, und vom sogenannten *asymmetrischen* Ramsey-Problem herrührt. Im Gegensatz zum obigen (symmetrischen) Problem versucht man im asymmetrischen Fall in jeder Farbe ein anderer Hypergraph zu vermeiden.

Wir betrachten auch sogenannte *online* Probleme. In den bisherigen Problemstellungen war der gesamte zu färbende Graph oder Hypergraph vor dem Färben bekannt. In online Spielen ist dies nicht mehr der Fall, und Knoten oder Kanten werden schrittweise in mehreren Runden enthüllt. Wir müssen die Färbung für die neu enthüllten Knoten oder Kanten sofort bestimmen, bevor die nächste Runde beginnt.

Wir betrachten das *balancierte Ramsey-Spiel* in dem der Spieler r Farben zur Verfügung hat, und in jeder Runde r Kanten eines anfänglich leeren Graphen auf n Knoten enthüllt werden. Der Spieler muss nach jeder Runde jeder Kante eine unterschiedliche Farbe zuweisen. Sein Ziel ist es eine einfarbige Kopie eines fixen Graphens F zu vermeiden. Wir vergleichen dieses Spiel mit dem ähnlichen *Achlioptas-Spiel*, in dem in jeder Runde auch r Kanten enthüllt werden, wo der Spieler aber $r-1$ davon wieder verwerfen darf. Das Ziel ist auch hier wieder eine Kopie von F zu verhindern. Das entspricht einem balancierten Ramsey-Spiel, in dem nur Kopien von F in einer fixen Farbe verboten sind. Gleich wie in den nicht-online (d.h. *offline*) Fällen hat die erwartete Dauer eines solchen Spiels einen Schwellenwert.

Krivelevich, Spöhel und Steger haben die Frage gestellt ob diese beiden online Spiele den gleichen Schwellenwert haben oder nicht. Die bisher bekannten Resultate zeigen nämlich, dass dies für alle Graphen (ausser Bäume) der Fall ist. Die Intuition hinter dieser Frage ist, dass das balancierte Ramsey-Spiel eigentlich nicht schwerer als das Achlioptas-Spiel sein könnte, weil der Spieler sowieso immer nur gegen eine schlechtestmögliche Farbe "ankämpft", und die restlichen $r-1$ Farben somit eigentlich irrelevant sind. Im letzten Teil dieser Arbeit beantworten wir diese Frage und zeigen, dass diese zwei Spiele auch für nicht-Bäume unterschiedliche Schwellenwerte haben können. Wir betrachten auch die Varianten dieser Spiele, in denen Knoten statt Kanten enthüllt und gefärbt werden. In diesem Fall gilt im Gegensatz zum Kantenfall, dass die beiden Spiele tatsächlich den gleichen Schwellenwert haben.

Acknowledgments

First and foremost I would like to thank my supervisor Angelika Steger for giving me the opportunity write this thesis and work in her group for the past years. I learned more than I would ever have thought possible, and I will always fondly remember the time spent there.

I am very grateful to Mathias Schacht for agreeing to co-referee this thesis.

My most sincere gratitude also goes to Reto Spöhel. He supervised my first struggles with Ramsey Theory during my master thesis, and his mentorship has left large footprints in my way of thinking about these problems.

My thanks also go to all my co-authors: Matthew Cook, Florian Jug, Christoph Krautz, Konstantinos Panagiotou, Yury Person, Ueli Peter, Reto Spöhel, Angelika Steger and Henning Thomas.

Of course thanks also go to the entire CSA family, not only for the scientific collaboration, but also for all the discussions during lunch, the coffee-breaks, Buchboden and much more. Roughly in order from "old hands" to "new recruits" these are: Dan, Florian, Christoph, Reto, Torsten, Thomas, Henning, Ueli, Johannes, Hafsteinn, Rajko, Andreas, Frank and Asaf.

I would also like to thank Thomas Rast. We met pretty much exactly 10 years ago, and our first conversation was about computers. Thankfully, it seems we never managed to stop. I am very grateful for his friendship.

Finally I would like to thank Ursi and my family. Mim, Pip, Laura, Carlo and Paolo: you are the best.

Chapter 1

Introduction

Ramsey theory is named after F. P. Ramsey and his seminal paper [Ram30]. Simplifying his results, he proved that for every integer $\ell \geq 1$ there exists a constant $R(\ell)$, such that no matter how we color the edges of a complete graph with $R(\ell)$ many vertices with the colors red or blue, we are always forced to create a monochromatic clique of size ℓ. The numbers $R(\ell)$ are called the Ramsey numbers.

This seemingly simple statement has deep implications and has sparked an entire branch of combinatorics called Ramsey theory. Loosely speaking its object of study are large structures and conditions under which coloring these structures ensures that certain other monochromatic substructures must always appear. For example in Ramsey's theorem we color large cliques and want to known conditions which guarantee that smaller monochromatic cliques appear. The answer in Ramsey's theorem is given in terms of the minimum number of vertices of the large clique, but we may also ask e.g. for the minimum number of edges in any graph which guarantees the same.

Ramsey-type questions are asked in many different settings (we describe some of them in the next chapter), but our interest in this thesis is in the setting of graphs and hypergraphs, and more specifically of *random* graphs and hypergraphs, which we introduce below. In this setting it is common to denote the property that every r-edge-coloring of some graph (or hypergraph) G contains a monochromatic copy of F by the notation $G \to (F)^e_r$, and $G \to (F)^v_r$ if instead of edges we color vertices. With this notation Ramsey's theorem may be stated as: for all $\ell \geq 1$ there exists some constant $R(\ell)$ such that $K_{R(\ell)} \to (K_\ell)^e_2$.

The name *random graph* refers to two different but closely related models which were introduced by Paul Erdős and Alfréd Rényi in 1959 [ER59] and independently in the same year by Edgar Gilbert [Gil59]. Erdős and Rényi first introduced the so called $G(n,m)$ model, in which a graph is chosen uniformly at random among all possible graphs on n vertices and m edges. The second model is called $G(n,p)$ and was introduced by Gilbert. In this second model we fix n vertices and each of the $\binom{n}{2}$ possible edges is present independently with probability p. In many situations, if $m = \binom{n}{2}p$ and $n^2p \to \infty$, these two models are equivalent.

The behavior of random graphs is usually studied for $n \to \infty$ and m or p are chosen as a function of n. Bollobás and Thomason [BT87] proved that every "nice" graph property, including many natural ones such as e.g. being connected or containing some small subgraph, have a so-called *threshold.* A threshold for a graph property $\mathcal{A}$ is some $m_0 = m_0(n)$ such that

$$\lim_{n\to\infty} \mathbb{P}[G(n,m) \text{ satisfies } \mathcal{A}] = \begin{cases} 0 & \text{if } m \ll m_0 \\ 1 & \text{if } m \gg m_0 \end{cases}$$

or, in the case of the $G(n,p)$ model, some $p_0 = p_0(n)$ such that

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \text{ satisfies } \mathcal{A}] = \begin{cases} 0 & \text{if } p \ll p_0 \\ 1 & \text{if } p \gg p_0. \end{cases}$$

(While of course any $p_0' = cp_0$ or $m_0' = cm_0$ for some positive constant c also satisfies the above equations, it is customary to talk about *the* threshold.)

A typical Ramsey-type question for random graphs is the one for the threshold for the property $G(n,p) \to (F)^e_r$ or $G(n,p) \to (F)^v_r$, where

F is some fixed graph. The study of this type of question was initiated by Łuczak, Ruciński and Voigt in [ŁRV92], where they considered vertex colorings and established the threshold for $G(n,p) \to (F)_2^v$ for arbitrary graphs F. They also considered edge colorings and established the threshold for $G(n,p) \to (K_3)_2^e$. In a subsequent series of papers Rödl and Ruciński [RR93, RR94, RR95] fully solved the edge coloring problem. They proved that for all $r \geq 2$ and all graphs F which are not forests of stars, or (in the case $r = 2$) forests of stars and at least one path on 3 edges, there exist two positive constants c, C such that

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \to (F)_r^e] = \begin{cases} 1 & \text{if } p \geq Cn^{-1/m_2(F)}, \\ 0 & \text{if } p \leq cn^{-1/m_2(F)}, \end{cases}$$

where $m_2(F) := \max_{H\subseteq F} \frac{e(H)-1}{v(H)-2}$. The cases of stars and paths of 3 edges are exceptional, and the corresponding thresholds are strictly lower than what would be obtained as above with $m_2(\cdot)$.

A *k-uniform hypergraph* is a straightforward generalization of the concept of a graph in which each edge contains not 2, but $k \geq 2$ many different vertices. A random k-uniform hypergraph $H^k(n,p)$ is accordingly obtained from an empty hypergraph on n vertices by selecting each of the $\binom{n}{k}$ possible edges with probability p independently. For $k = 2$ this is equivalent to the $G(n,p)$ model.

Rödl and Ruciński [RR98] first considered Ramsey-type properties for random hypergraphs, specifically the question for the threshold of $H^k(n,p) \to (F)_r^e$. They conjectured that at least the upper bound in their results for the graph case should transfer to the case of hypergraphs by replacing $m_2(\cdot)$ by $m_k(F) := \max_{H\subseteq F} \frac{e(H)-1}{v(H)-k}$. They verified this in the case of two colors and the complete 3-uniform hypergraph on 4 vertices. Together with Schacht in 2007 they extended this to include all k-partite, k-uniform hypergraphs. Recently Friedgut, Rödl and Schacht [FRS10], and independently Conlon and Gowers [CG10], proved the full conjecture. It is widely believed that this upper bound is tight for most hypergraphs F, however no corresponding lower bounds are known. In this thesis we address this question and show a matching lower bound for a large class of hypergraphs. We also present examples of hypergraphs for which the upper bound is not tight.

A related Ramsey-type question is the so-called *asymmetric* Ramsey problem in random graphs. Instead of avoiding a monochromatic copy of the same graph F in all colors as above, in thea symmetric Ramsey case we want to avoid a graph F_1 in red, a graph F_2 in blue, and so on for all $r \geq 2$ colors. This problem was first introduced by Kohayakawa and Kreuter in [KK97], where they determined the threshold for the appearance of some combinations of cycles (and some more general cases). The authors also conjectured a general threshold. Kohayakawa, Schacht and Spöhel [KSS] proved an upper bound for $r = 2$ and two graphs G and H satisfying some mild conditions. Nenadov [Nen13] recently generalized their results to random k-uniform hypergraphs.

In all settings discussed so far the entire instance of $G(n, p)$ is always known before coloring its edges. In so-called *online* problems this is no longer the case. In such settings edges of a random graph on n vertices are revealed a little at a time and we must immediately commit to a color before the next ones are revealed. The goal is to avoid creating some monochromatic structure, e.g. a monochromatic copy of some fixed graph F. These questions are usually stated as a *game* in which a player called Painter colors the edges or vertices. The reason for this is that the typical duration of such a game depends heavily on the *strategy* of Painter. For "reasonable" strategies the typical duration of the game exhibits a threshold type behavior, i.e. there exists some $N_0 = N_0(n)$ such that Painter can win with probability $1 - o(1)$ using this strategy if the game is played for $m \ll N_0$ many rounds, while she loses it with probability $1 - o(1)$ for $m \gg N_0$. If we ask for the threshold for a given game, we instead ask for a threshold such that below it there exists a strategy such that Painter can win with high probability, while above it she almost surely loses, regardless of her strategy.

This type of question was first asked by Friedgut, Kohayakawa, Rödl, Ruciński and Prasad in [FKR$^+$03], where they established a threshold of $n^{4/3}$ for the typical duration of the game in which one edge is revealed at a time which must immediately be colored either red or blue while avoiding the creation of a monochromatic triangle.

Of interest in this thesis is in particular the *balanced online Ramsey game*. In this game Painter has r colors at her disposal and r edges are revealed at a time. In each round Painter must assign a different color to each edge and her goal is to avoid creating a monochromatic copy of some fixed graph F. The came is called balanced because its rules force all color classes to have exactly the same size. This game was first introduced by Marciniszyn, Mitsche and Stojaković [MMS07], which found the threshold for avoiding cycles with 2 colors. Prakash, Spöhel and Thomas [PST09] generalized these results to an arbitrary number of colors and proved thresholds valid for a large class of graphs including cycles and cliques. They also studied the vertex variant of this game and proved similar results.

A related type of problem are so-called Achlioptas processes. These have their origin in a question asked by Achlioptas: if we start with an empty graph on n vertices, reveal 2 edges at a time, and must immediately choose which one to keep and which one to discard, is there a way to select the edges such that we can accelerate or delay the appearance of certain substructures in this graph graph? His original question concerned delaying the appearance of a so-called *giant component*, i.e. a connected component of size linear in the number of vertices, and this was answered positively by Bohman and Frieze in [BF01]. This result spurred a lot of interest in this kind of processes, and they were generalized in many ways. One example of this are the results by Krivelevich, Loh and Sudakov [KLS09], which considered the F-avoidance game in the Achlioptas setting. This process may be seen as a weaker version of the balanced Ramsey game in which one color represents chosen edges and all others discarded ones. Accordingly the F avoidance game in the Achlioptas process is essentially a balanced Ramsey game for which only red copies of F are dangerous.

1.1 Results in this thesis

1.1.1 A Randomized Version of Ramsey's Theorem

The randomization of Ramsey's theorem by Rödl and Ruciński [RR93, RR95] asks for a threshold for p and the property $G(n,p) \to (F)^e_r$. Their randomization can also be interpreted as the question of how many edges one needs to remove from the complete graph on n vertices, such that an r-edge-coloring without a monochromatic copy of F becomes possible.

In Chapter 4 we consider a different randomization that was recently suggested by Allen, Böttcher, Hladký and Piguet [ABHP13]. Let $\mathcal{R}_F(n,q)$ be a random subset of all copies of F in K_n, in which every copy is present independently with probability q. For which functions $q = q(n)$ is it possible to color the edges of K_n with r colors such that no monochromatic copy of F is contained in $\mathcal{R}_F(n,q)$? In other words they suggest that instead of removing edges we remove copies of F from the set of those that cause us to lose if they are monochromatic.

We answer this question for strictly 2-balanced graphs F. Moreover, we combine both randomizations and prove a threshold result for the property that there exists an r-edge-coloring of $G(n,p)$ such that no monochromatic copy of F is contained in $\mathcal{R}_F(n,q)$.

These results are joint work with Yury Person, Angelika Steger and Henning Thomas [GPST12].

1.1.2 Lower bounds for Ramsey properties of random hypergraphs

In Chapter 5 we consider the extension of the results by Rödl and Ruciński to the hypergraph setting. It is widely believed that the upper bound by Friedgut, Rödl and Schacht (and independently Conlon and Gowers) discussed above is tight in most cases. So far however no explicit results are known. We prove a matching lower bound which holds for all hypergraphs F which do not have an "obvious counterexample" of finite size. In the graph case forests of stars and forests of stars and at least one path on 3 edges are in a certain sense exceptional. We show that in the hypergraph case there are many more exceptional cases in which the upper bound is not tight. We also show that in contrast to the graph case in at least one example the threshold for the symmetric problem is given by an instance of the asymmetric Ramsey problem.

These results are joint work with Yury Person, Angelika Steger and Henning Thomas.

1.1.3 Balanced Coloring Games in Random Graphs

In Chapter 6 we turn our attention to online games. We compare the balanced Ramsey game with a game similar to the Achlioptas process outlined above. In this Achlioptas game r instead of 2 edges are revealed at a time, and instead of choosing one edge and discarding the remaining $r-1$, we have r colors and must assign a different one to each revealed edge. The goal is to avoid creating a red copy of some fixed graph F for as long as possible. In other words this game is a weaker version of the balanced game in which only one color is dangerous.

Krivelevich, Spöhel and Steger noted in [KSS10] that all known results for graphs F which are not a forest show that the threshold for these two games coincide. They ask whether this is true for all non-forest F.

We answer this question negatively for the edge version of both games. We show that there is an infinite family of non-forests F for which the balanced Ramsey game has a different threshold than the Achlioptas game. We also consider the natural vertex analogues of both games and show that their thresholds coincide for all graphs F, in contrast to our results for the edge case.

These results are joint work with Reto Spöhel [GS14].

Chapter 2

Notation

In this chapter we introduce notation that is used throughout the thesis. Most chapter-specific notation is instead introduced in the chapters themselves.

2.1 Graphs and hypergraphs

Unless specified otherwise all graphs or hypergraphs in this thesis are simple, i.e. undirected and without multiple edges. All hypergraphs are k-*uniform* for some integer $k \geq 2$, in other words such that the number of vertices contained in each hyperedge is equal to k. We assume that every graph or hypergraph contains at least one vertex and we say that it is *empty* if it contains no edge, and *non-empty* otherwise. We usually use the term edge instead of hyperedge for hypergraphs as well, unless (as above) we wish to emphasize that we are talking about an edge of a hypergraph.

For a graph or hypergraph G use the notation $V(G)$ and $E(G)$ for the vertex and (hyper)edge-set respectively, and $v(G)$, $e(G)$ for their respective sizes. The degree of a vertex v in G is indicated by $\deg_G(v)$ or just $\deg(v)$ if G is implied from the context. The minimum degree over all vertices is $\delta(G)$, and the maximum is $\Delta(G)$.

We often discuss graphs and hypergraphs side by side and compare results for graphs to their counterpart for hypergraphs. Strictly speaking for $k = 2$ these are the same, so in these contexts "hypergraph" implicitly refers to the case of uniformity $k \geq 3$.

Throughout the thesis K_ℓ denotes a complete graph on ℓ vertices, C_ℓ a cycle on ℓ edges and P_ℓ a path on ℓ edges. Further S_ℓ is a star with ℓ rays, and $K_{s,t}$ is the complete bipartite graph with partitions of s and t vertices. The complete k-uniform hypergraph on $\ell \geq k$ vertices is denoted by K_ℓ^k. A *matching* is a graph with maximum degree 1.

The subgraph (or subhypergraph) of F induced by a set $W \subseteq V(F)$ is $F[W]$. If W is not a subset of $V(F)$, then $F[W]$ denotes $F[W \cap V(F)]$. By $F \uplus G$ we denote the disjoint union of two graphs or hypergraphs F and G.

In Chapter 6 we use the notions of ordered and of r-matched graphs (which are explained there). Where notions such as the vertex set, edge set, degree etc. are transferable to these types of graphs we use the same notation as for simple graphs.

In the context of Ramsey-type problems we use the notation $G \to (F)_r^e$ to indicate that every edge coloring of a graph G with r colors creates a monochromatic copy of a graph F. We use $G \to (F)_r^v$ if instead we color vertices. Both G and F may also be hypergraphs.

With $G(n,p)$ we always indicate a random graph on n vertices in which every possible edge appears with probability p independently. With $G(n,m)$ a random graph selected uniformly at random among all graphs with n vertices and m edges. For $p = m = 1$ this notation may be ambiguous, but the risk of confusion is minimal in our settings. For hypergraphs we use $H^k(n,p)$ to denote the random k-uniform hypergraph on n vertices in which every possible k-tuple of distinct vertices is a hyperedge with probability p independently.

2.2 Asymptotics

Unless stated otherwise all our asymptotics are with respect to n, which usually indicates the number of vertices of some random graph or hypergraph. We use the standard Landau symbols $\omega, \Omega, o, \mathcal{O}, \Theta$ with their usual meaning, and also use $f(n) \ll g(n)$ for $f(n) = o(g(n))$ and $f(n) \gg g(n)$ for $f(n) = \omega(g(n))$.

In the context of probabilities we use the phrase *asymptotically almost surely*, (a.a.s.) meaning that for a series of events $\mathcal{E}_n$ it holds that $\mathbb{P}[\mathcal{E}_n] = 1 - o(1)$ as $n \to \infty$. We may also use *with high probability* (w.h.p.) to mean the same.

2.3 Useful tools

On several occasions we make use use the following proposition.

Proposition 2.1. *Let $a, c, \beta \geq 0$ and $b, d > 0$. Then we have*

$$\begin{aligned} \frac{a}{b} > \beta \quad \text{and} \quad \frac{c}{d} \geq \beta \quad &\Longrightarrow \quad \frac{a+c}{b+d} > \beta, \text{ and} \\ \frac{a}{b} < \beta \quad \text{and} \quad \frac{c}{d} \leq \beta \quad &\Longrightarrow \quad \frac{a+c}{b+d} < \beta. \end{aligned}$$

Further, if $b \neq d$, we have

$$\frac{a-c}{b-d} \gtrless \frac{a}{b} \Longleftrightarrow \frac{a}{b} \gtrless \frac{c}{d} \quad \text{and} \quad \frac{a+c}{b+d} \gtrless \frac{a}{b} \Longleftrightarrow \frac{a}{b} \lessgtr \frac{c}{d}.$$

Chapter 3

Background and related work

3.1 Ramsey theory

Ramsey theory has its roots in a statement proved by Frank P. Ramsey [Ram30] in 1930. On a very high level Ramsey theory studies conditions which ensure that order appears when partitioning large structures.

Before we present Ramsey's theorem consider this simpler observation, which has come to be known as the pigeonhole principle. It states that if n pigeons are sitting in m pigeonholes, and $n > m$, then there must exist some pigeonhole containing two or more pigeons. This observation is trivial, but it contains many of the features of Ramsey theory. We partition pigeons into pigeonholes, the "order" in this case are shared pigeonholes, and the necessary condition for it to appear is $n > m$.

Theorem 3.1 (Ramsey Theorem, [Ram30]). *For all positive integers k, ℓ, r there exists a smallest integer $n(k, \ell, r)$ such that if the set of subsets of size k of a set N of size $n \geq n(k, \ell, r)$ is colored arbitrarily with r colors, then there exists some subset $N' \subseteq N$ of size ℓ such that all the subsets of size k of N' have the same color.*

The above is the finite version of Ramsey's theorem, in his seminal work he also proved a stronger version which holds for countably infinite sets.

Ramsey's theorem is most commonly presented as a much more accessible statement about complete graphs K_n on n vertices. We obtain it from the theorem above by choosing N as the set of vertices of a K_n. Then we set $k = 2$, i.e. we consider the set of all 2-element subsets of N, or equivalently consider the edge set of K_n. The theorem above then guarantees the existence of a smaller complete graph K_ℓ which is monochromatic, regardless of how we color the edges of K_n.

Theorem 3.2 (Ramsey Theorem for graphs). *For any $\ell > 0$ there exists a number $R(\ell)$ such that regardless of how one colors the edges of the complete graph K_n on $n \geq R(\ell)$ vertices with 2 colors, it always contains a monochromatic complete graph on ℓ vertices.*

The numbers $R(\ell)$ are called Ramsey numbers. It is known that $R(1) = 1$, $R(2) = 2$, $R(3) = 6$ and $R(4) = 18$. For $R(5)$ the best known bounds are $43 \leq R(5) \leq 49$, [Exo89, MR97]. Any higher numbers are similarly unknown, an extensive survey of known bounds is found at [Rad11]. These numbers are famously difficult to compute. Graham and Spencer [GS90] recount the following story, attributed to Paul Erdős.

> Aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find $R(5)$. We could marshal the world's best minds and fastest computers, and within the year we could probably calculate the value. If the aliens demand $R(6)$ we would have no choice but to launch a preemptive attack.

The asymptotic behavior of $R(\ell)$ as $\ell \to \infty$ is one of the biggest open questions in Ramsey theory. With comparatively simple arguments Erdős and Szekeres [ES35, Erd47] proved that

$$2^{\ell/2} \leq R(\ell) \leq 2^{2\ell}.$$

Despite considerable effort the constants 1/2 and 2 in the exponent are still best possible. The results by Erdős and Szekeres are actually slightly tighter, in the above we omitted some small factors. There have been several other improvements over the years [Spe75, GR87, Tho88, Con09], however the two constants in the exponents remained unchanged. Even the most minute improvement to them would be considered a breakthrough.

Ramsey numbers are generalized in several ways. Of considerable interest are the *asymmetric* Ramsey numbers, denoted $R(\ell, k)$. Here we ask for the smallest integer $R(\ell, k)$ so that all 2 colorings of the edges of a complete graph on $R(\ell, k)$ many vertices contains either a red K_ℓ or a blue K_k.

Another generalization is to replace the complete graphs by arbitrary graphs F and G. We write $G \to (F)^e_r$ if every r coloring of the edges of G induces a monochromatic copy of F, and we write $G \to (F)^v_r$ for the same statement if we color vertices instead. An example of such a statement would be $K_{3,7} \to (C_4)^e_2$, which says that every 2-edge-coloring of a complete bipartite graph with partitions of 3 and 7 vertices must contain a monochromatic cycle of length 4.

A further related topic are the so-called *size-Ramsey* numbers $r_e(F)$ for a given fixed graph F. $r_e(F)$ is defined as the smallest number of edges in any graph G satisfying $G \to (F)^e_2$. The study of this type of question was first proposed by Erdős, Faudree, Rousseau and Schelp [EFRS78], which among other things proved that $r_e(K_\ell) = \binom{R(K_\ell)}{2}$.

Ramsey theory is however not only confined to graphs. One of its earliest results is by Schur [Sch16] and it even predates Ramsey's famous paper. It states that for some fixed number of colors r there exists some integer $n_0 = n_0(r)$ such that every r-coloring of the integers $1, \ldots, n_0$ contains a monochromatic solution to the equation $x + y = z$. Van der Waerden [vdW27] proved a similar theorem concerning the appearance of a monochromatic arithmetic progression. These results were generalized by Rado [Rad33] to solutions of systems of linear equations. Many other profound theorems stem from here, such as Szemerédi's theorem on arithmetic progressions [Sze75] or the Hales–Jewett theorem [HJ63]. An overview over many of these results can be found in [GRS90].

3.2 Random graphs

As discussed in the introduction the concept of a random graph usually refers to two similar models. For the $G(n, p)$ model, first proposed by Gilbert [Gil59], we start with an empty graph on n vertices and select each of the $\binom{n}{2}$ possible edges independently at random with probability p. The $G(n, m)$ model on the other hand was introduced by Erdős and Rényi [ER59]. A $G(n, m)$ random graph is chosen uniformly among all graphs with n vertices and m edges.

Glimpses of these ideas are already present in the 1947 paper by Erdős [Erd47] in which he established his lower bound for the asymptotic behavior of $R(\ell)$ we discussed above. This paper uses the uniform distribution on all graphs on n vertices, which is given by $G(n, p)$ for $p = 1/2$. It also represents one of the first applications of the probabilistic method.

It is undoubtedly due to Erdős and Rényi's impressive work in the 1960s [ER61a, ER61b, ER63, ER64, ER66, ER68] that random graph theory enjoys the attention it has today. A comprehensive introduction to the history of random graphs can be found in [KR97].

As we shall see these models are essentially equivalent in many respects, but depending on the problem under consideration one or the other may seem more natural. The $G(n, m)$ model is especially amenable to questions concerning the *evolution* of random graphs. For example, if we start with an empty graph on n vertices and add one random edge after the other, how many edges do we need to add in order for the graph to be connected? After m steps of this *random graph process* the graph is distributed like $G(n, m)$ and it is natural to ask for an answer in terms of m. In his 1959 paper Gilbert considered the same question, but his take was a completely different. He considered the example of n central offices of a telephone network (he was a researcher at Bell Laboratories after all), such that the probability that there is an idle direct line between any two offices is p. Then he asked for values of p such that every office is able to call everyone else, routing the call via other offices if necessary. This is equivalent to asking for which values of p the random graph $G(n, p)$ is with high probability connected.

Recall that a *threshold* for a graph property $\mathcal{A}$ is some $p_0 = p_0(n)$ such that

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \text{ satisfies } \mathcal{A}] = \begin{cases} 0 & \text{if } p \ll p_0 \\ 1 & \text{if } p \gg p_0. \end{cases}$$

(The definition for the $G(n,m)$ model is equivalent.)

It is easy to see that if $m = \binom{n}{2}p$ and $n^2 p \to \infty$, then $G(n,m)$ and $G(n,p)$ have "about the same" number of edges, so one would expect them to behave similarly. This turns out to be true for all *monotone* graph properties, i.e. properties which if true for a graph G remain true if we add more edges (but not vertices) to G. Bollobás and Thomason [BT87] proved that every monotone graph property has a threshold. Further, if m_0 is the threshold for a monotone property in $G(n,m)$, then $p_0 = m_0/n^2$ is the threshold for the same property in $G(n,p)$. In other words, if we are studying a monotone property, we may switch between the two models almost at will. This is particularly convenient, as in the $G(n,p)$ model all edges appear independently of each other, which in many cases makes for simpler proofs.

Many interesting graph properties such as e.g. connectedness, the existence of a component of size linear in the number of vertices, the existence of a fixed subgraph etc., are actually monotone. For this reason a large fraction of the results in random graph theory are about finding exact expressions for the thresholds of some monotone property.

One of the most useful results for working with random graphs is arguably Bollobás result [Bol81] on the small subgraph containment problem. It establishes the threshold for the property that $G(n,p)$ contains a copy of some fixed graph F.

Theorem 3.3 ([Bol81])**.** *Let F be a non-empty graph, and set*

$$m(F) := \max_{\substack{H \subseteq F \\ v(H) \geq 1}} d(H) \quad \textit{where} \quad d(H) := \frac{e(H)}{v(H)}.$$

Then

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \textit{ contains a copy of } F] = \begin{cases} 0 & \textit{if } p \ll n^{-1/m(F)}, \\ 1 & \textit{if } p \gg n^{-1/m(F)}. \end{cases}$$

We call graphs F for which $m(F) = d(F)$ *balanced*, and *strictly balanced* if $m(F) > d(H)$ for all non-empty $H \subsetneq F$.

The above result does not hold for $p = \Theta(n^{-1/m(F)})$. Bollobás [Bol81] and independently Karoński and Ruciński [KR83] proved that in this situation the probability that $G(n,p)$ contains F is bounded away from 0 or 1.

Theorem 3.4 ([Bol81, KR83]). *Let F be a strictly balanced graph and $c > 0$ some constant. If $np^{m(F)} \to c$, then the number of copies of F in $G(n,p)$ converges in distribution to a Poisson random variable with expectation $\lambda = c^{v(G)}/\mathrm{Aut}(F)$, where $\mathrm{Aut}(F)$ is the set of automorphisms of F.*

3.3 Ramsey theory for random graphs

The intersection of Ramsey theory and the theory of random graphs was first explored by Łuczak, Ruciński and Voigt in [ŁRV92], where they considered vertex colorings and established the threshold for $G(n,p) \to (F)_2^v$ for arbitrary graphs F. They also considered edge colorings and established the threshold for $G(n,p) \to (K_3)_2^e$, i.e. the case of triangles and 2 colors.

The edge coloring problem proved more difficult, but in a subsequent series of papers Rödl and Ruciński first established a lower bound for the threshold [RR93], then extended their result for triangles to an arbitrary number of colors [RR94] and finally proved the threshold for $G(n,p) \to (F)_r^e$ in [RR95] for (almost) all graphs F. As pointed out by Friedgut and Krivelevich [FK00] they missed that the corner case of 2 colors and F a path of 3 edges is exceptional.

To state Rödl and Ruciński's result result we first need the following definition. Let F be a graph, set

$$d_2(F) := \begin{cases} 0 & \text{if } e(F) = 0, \\ 1/2 & \text{if } F = K_2, \\ \frac{e(F)-1}{v(F)-2} & \text{otherwise,} \end{cases} \qquad \text{and} \qquad m_2(F) := \max_{H \subseteq F} d_2(H).$$

We call $m_2(F)$ the 2-density of F, and say that F is 2-balanced if $d_2(F) = m_2(F)$ and strictly 2-balanced if $m_2(F) > d_2(H)$ for all $H \subsetneq F$.

Theorem 3.5 ([RR93],[RR95],[FK00]). *Let $r \geq 2$ be an integer and let F be a graph with at least one edge and which is not a forest of stars or, in the case $r = 2$, not a forest of stars and at least one path of 3 edges. Then there exist positive constants $c = c(F,r)$ and $C = C(F,r)$ such that*

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \to (F)_r^e] = \begin{cases} 0 & \text{if } p \leq cn^{-1/m_2(F)}, \\ 1 & \text{if } p \geq Cn^{-1/m_2(F)}. \end{cases}$$

There is a simple intuition behind the bound in the above theorem. For 2-balanced graphs it holds that for $p = c'n^{-1/m_2(F)}$, where $c' > 0$ is some constant, the number of copies of F and the number of edges in $G(n,p)$ have both the same order of magnitude. This implies that the number of copies of F on each edge is directly proportional to c'. For small values of c' (i.e. smaller than c) the number of copies of F per edge is small, and we do not expect copies of F to overlap much. Intuitively this should allow us to color $G(n,p)$ without a monochromatic copy of F. If on the other hand c' is very large then we expect every edge to be contained in many copies of F which all overlap heavily, and this makes finding a coloring without monochromatic F impossible.

Stars and paths of length 3 are excluded from the theorem, because they have a different threshold. For a star S_k with k edges it holds by the pigeon hole principle that the star with $r(k-1)+1$ many edges cannot be r-colored without creating a monochromatic S_k. By Theorem 3.3 the threshold for the appearance of the $(r(k-1)+1)$-star is lower than $m_2(S_k)$. For the path P_3 with 3 edges a similar construction is possible: every odd cycle of length 5 or more with one additional pending edge attached to each vertex (we call this a sunshine graph) cannot be 2-edge-colored without a monochromatic P_3. The threshold for the appearance of any such sunshine graph is asymptotically equal to $m_2(P_3)$. By Theorem 3.4 however a 0-statement as in the theorem above cannot hold. For any $c' > 0$ and any $p = c'n^{-1/m_2(F)}$ the probability that a sunshine graph appears in $G(n,p)$ is bounded away from 0.

A natural question which arises in the context of Theorem 3.5 is the one for precise values for the constants c and C. Friedgut and Krivelevich [FK00] proved that in the case of trees which are not stars or paths of length 3 they are actually equal, and that the property $G(n,p) \to (T)^e_r$ has a so-called *sharp* threshold. A caveat is that this is a purely existential result, and does not give a value for c. Further it may be possible that the value of this constant fluctuates with n, and it is not even known whether it converges in the limit $n \to \infty$. Firedgut, Rödl, Ruciński and Tetali [FRRT06] later showed that this also holds in the case of triangles.

In [FK00] Friedgut and Krivelevich also showed that similar results hold as well for the vertex version of Theorem 3.5 and a large class of graphs including cliques.

The randomization in Theorem 3.5 may be seen as choosing the number of edges of a complete graph K_n which we need to randomly remove such that we can color the remaining ones with r colors without creating a monochromatic copy of a forbidden graph F. A different type of randomization was recently suggested by Allen, Böttcher, Hladký and Piguet [ABHP13]. They suggest that we may leave all edges in the complete graph K_n, but instead for each possible copy of F in K_n choose whether it is forbidden or not with probability q independently at random. The question is then for a threshold for q such that we can color the edges of K_n with r colors such that no monochromatic copy of F is among the set of those which are forbidden. Allen et al. considered this type of randomization applied to Turán's theorem. We investigate their question for the Ramsey problem in Chapter 4, and also combine both types of randomization.

Another problem in the intersection of Ramsey theory and random graphs is the so-called *asymmetric Ramsey problem in random graphs*. Instead of avoiding a monochromatic copy of the same graph F in all colors as above, in the asymmetric Ramsey case we want to avoid a graph F_1 in red, a graph F_2 in blue, and so on for all $r \geq 2$ colors. Similarly to the classical Ramsey case, we denote the fact that all edge colorings of a graph G contain at least one monochromatic copy of F_i in its respective color as $G \rightarrow (F_1, \dots, F_r)$. If all F_i are equal this reduces to the (symmetric) Ramsey case.

This problem was first introduced by Kreuter [Kre96] for the vertex coloring case and by Kohayakawa and Kreuter in [KK97] for the edge case. For the vertex case Kreuter proved a result which holds for almost every graph. For the edge case Kohayakawa and Kreuter determined the threshold for the appearance of some combinations of cycles (and some more general cases). They also conjectured that in the general case the threshold is determined by the function below in the same sense that $m_2(\cdot)$ determines the threshold for the symmetric Ramsey problem in Theorem 3.5.

Let G_1, G_2 be two graphs with at least one edge and such that $m_2(G_1) \geq m_2(G_2)$. Then define

$$m_2(G_1, G_2) = \max\Big\{\frac{e(G_1')}{v(G_1') - 2 + 1/m_2(G_2)} \mid G_1' \subseteq G_1, e(G_1') \geq 1\Big\}.$$

Note that if $G_1 = G_2$, then $m_k(G_1, G_2) = m_k(G_1)$. We say that G_1 is strictly balanced with respect to $m_k(\cdot, G_2)$ if no strict subgraph $G_1' \subsetneq G_1$ with at least one edge maximizes the above equation.

While the above function seems only to apply to the case $r = 2$, it turns out that even for larger values of r the two graphs with maximal 2-density fully determine the threshold. Progress towards proving the conjecture was made by Marciniszyn, Skokan, Spöhel and Steger in [MSSS09], where they proved the 0-statement for cliques. Kohayakawa, Schacht and Spöhel [KSS] then proved the upper bound for $r = 2$ and two graphs G and H satisfying some mild conditions.

In [MSSS09] Marciniszyn et al. expanded on the results by Kohayakawa et al. [KK97] and also proved the general version of the 1-statement, provided that the so-called KŁR-conjecture [KŁR97] holds. At the time of their writing this conjecture was still open. Recently Balogh, Morris and Samotij [BMS12a], and independently Saxton and Thomason [ST12], proved the KŁR-conjecture. From their results a 1-statement for r many 2-balanced graphs, where the one of maximal 2-density has to be strictly 2-balanced, follows.

3.4 Ramsey theory for random hypergraphs

A natural generalization of the Ramsey-type results in random graphs from the previous section consists of transferring them to the setting of random k-uniform hypergraphs. A random k-uniform hypergraph on n vertices $H^k(n,p)$ is such that each possible subset of k vertices forms an edge with probability p independently.

Rödl and Ruciński initiated the study of this type of question in [RR98], where they conjectured that for hypergraphs the same intuition must hold as for the graph case, and that a monochromatic copy of F appears in every coloring whenever the expected number of copies of F per hyperedge exceeds a large constant. This suggests the following generalization from the graph case. Define for a k-uniform hypergraph F

$$d_k(F) := \begin{cases} 0 & \text{if } e(F) = 0, \\ 1/k & \text{if } e(F) = 1, v(F) = k, \\ \frac{e(F)-1}{v(F)-k} & \text{otherwise,} \end{cases} \quad \text{and} \quad m_k(F) := \max_{H \subseteq F} d_k(H).$$

As in the graph case we call $m_k(F)$ the k-density of F, and say that F is k-balanced if $d_k(F) = m_k(F)$ and strictly k-balanced if $m_k(F) > d_k(H)$ for all $H \subseteq F$.

The conjecture is that $m_k(\cdot)$ determines an upper bound on the threshold for the hypergraph case in the same way as $m_2(\cdot)$ does for the graph case. They proved this for the 3-uniform clique on 4 vertices and 2 colors.

Together with Schacht they later proved their conjecture for k-partite k-uniform hypergraphs [RRS07]. Recently Friedgut, Rödl and Schacht [FRS10] proved the conjecture for arbitrary k-uniform hypergraphs. Similar results were obtained independently by Conlon and Gowers [CG10].

Theorem 3.6 ([FRS10], [CG10])**.** *Let F be a k-uniform hypergraph with maximum degree at least 2, and let $r \geq 2$. There exists a constant $C > 0$ such that for $p = p(n)$ satisfying $p \geq Cn^{-1/m_k(F)}$ we have*

$$\lim_{n\to\infty} \mathbb{P}[H^k(n,p) \to (F)_r^e] = 1.$$

The above bound is widely believed to be tight in most cases, with some exceptions similar to those in Theorem 3.5. In Chapter 5 we confirm this for a large class of hypergraphs. However we also show that there are many examples for which the above bound is not tight, including one for which the bound is in fact given by the asymmetric Ramsey problem for hypergraphs.

In this context, Nenadov [Nen13] recently extended the graph-case result by Kohayakawa et al. in [KSS] to k-uniform hypergraphs, exploiting a containment theorem by Saxton and Thomason [ST12].

Theorem 3.7 ([Nen13])**.** *Let $G_1, \dots, G_r$ be k-uniform hypergraphs such that $m_k(G_1) \geq m_k(G_2) \geq \cdots \geq m_k(G_r)$ and such that G_1 is strictly balanced with respect to $m_k(\cdot, G_2)$. Then there exists a constant $C > 0$ such that for $p = p(n) \geq Cn^{-1/m_k(G_1,G_2)}$*

$$\lim_{n\to\infty} \mathbb{P}[H^k(n,p) \to (G_1, \dots, G_r)] = 1.$$

Here again $m_k(\cdot,\cdot)$ generalizes the quantity $m_2(\cdot,\cdot)$ from the graph case, and for k-uniform hypergraphs G_1, G_2 with at least one edge and such that $m_k(G_1) \geq m_k(G_2)$ is defined as follows.

$$m_k(G_1, G_2) = \max\Big\{\frac{e(G_1')}{v(G_1') - k + 1/m_k(G_2)} \mid G_1' \subseteq G_1, e(G_1') \geq 1\Big\}.$$

We say that G_1 is strictly balanced with respect to $m_k(\cdot, G_2)$ if it holds that $m_k(G_1, G_2) > m_k(H_1, G_2)$ for all strict subgraphs $H_1 \subsetneq G_1$.

3.5 Online Ramsey games in random graphs

All the results in the previous two sections implicitly presuppose that we can look at the entire random graph or hypergraph before committing to a coloring of the edges or vertices.

A different type of question are so-called *online* problems, in which information is revealed over time, and we do not have complete information before deciding how to color an edge or vertex. To avoid confusion, in this setting we refer to the results with complete information as *offline* results.

An example of such an online problem is the F avoidance game for some graph F. The game starts with an empty graph on n vertices and in each round one random edge is revealed. The player, we call her Painter, has r colors and must immediately decide which one to assign to the new edge. Her goal is to avoid a monochromatic copy of F for as long as possible. This can be seen as an online version of the offline problem of avoiding a monochromatic copy of F in the random graph in Theorem 3.5.

Note that we introduced the above as a *game* with one player, and not in the more abstract setting we used for the offline results in the previous sections. The reason is that in this setting the question of how long Painter can play is dependent on *how* exactly she plays. For any given strategy there exists some threshold N_0 (as for the offline case for reasonable graph properties and strategies this is always the case, cf. [MSS09a]) for the number of edges such that above this threshold Painter with high probability loses with that strategy.

Consider the triangle avoidance game, i.e. the case $F = K_3$. After m rounds the game board is distributed as $G(n, m)$, so for example if Painter decides to always use only one color, then she loses as soon as a triangle appears in $G(n, m)$, which is the case for $m \gg n^{-1}$.

An obvious upper bound for the *length* of the triangle avoidance game (i.e. the number of rounds that are typically played) is given by the offline result in Theorem 3.5. Regardless of how one colors the edges of $G(n, m)$ with two colors, after more than $Cn^{3/2}$ rounds the game is over with probability $1 - o(1)$.

The question one seeks to answer in this type of problem is for a *threshold* $N_0(r, n)$ such that for $m \ll N_0(r, n)$ there exists a strategy for Painter which allows her to play with high probability for at least m steps, while for $m \gg N_0(r, n)$ she loses with probability $1 - o(1)$ regardless of her strategy. Note that in contrast to the threshold of the offline case, the order of magnitude of online thresholds usually depends on the number of colors. It is also often observed (informally) that the threshold formula for $r \to \infty$ matches the corresponding threshold for the offline problem.

The study of this type of problem was initiated by Fiedgut, Kohayakawa, Rödl, Ruciński and Tetali [FKR+03]. They proved that for 2 colors there exists a strategy for the triangle avoidance game such that the game lasts $\Theta(n^{4/3})$ rounds, and they also showed that no strategy can do better than that.

Marciniszyn, Spöhel and Steger [MSS05] then proved the threshold for $r = 2$ and F any clique. Later the same authors [MSS09a] proved lower bounds for the threshold for a large class of graphs which includes cycles and cliques and for an arbitrary (but fixed) number of colors. In [MSS09b] they proved matching upper bounds for the case $r = 2$. They conjectured that their lower bounds are tight even for 3 or more colors, however the best upper bound in the case $r \geq 3$ remained those given by the offline setting.

Belfrage, Mütze and Spöhel [BMS12b] proved that this probabilistic one player game can be linked to a deterministic 2-player game which they call the *Builder-Painter game*. In this setting Builder chooses which edges to present to Painter, and Painter tries to color them while avoiding a monochromatic copy of some graph F. The authors show that if Builder has a winning strategy which only creates graphs G with $m(G) \leq d$, then $n^{2-1/d}$ is a threshold for the original 1 player game. The intuition behind this result is that in a random graph $G(n, m)$, for $m \gg n^{2-1/d}$ the number of copies of *every* such G is $\omega(1)$. In particular this implies that with high probability the edges of at least one of these graphs G is presented to Painter in the order in which Builder would play, and thus causes Painter to lose.

Balogh and Butterfield [BB10] used this Builder-Painter game to prove for the triangle avoidance game and $r = 3$ that there is some constant c' such that $n^{3/2-c'}$ is an upper bound for the threshold of the online game. The corresponding offline bound is at $\Theta(n^{3/2})$, and this was the first result showing that the thresholds for the offline and online games differ for $r \geq 3$. This threshold however is far from the lower bound in [MSS09a]. Noever [Noe12] later proved an upper bound for the online triangle avoidance game and arbitrary number of colors which matches the lower bound in [MSS09a] and confirms their conjecture for triangles.

The vertex-coloring game corresponding to this was studied by Marciniszyn and Spöhel [MS07], which established thresholds for a large class of graph which includes cliques and cycles. Mütze, Rast and Spöhel [MRS11] then solved the problem in full generality.

3.6 The Achlioptas process and the balanced Ramsey game

A different kind of online games have their origin in a question posed by Dimitris Achlioptas: if in the random graph process we reveal 2 edges at a time and we must immediately choose which one to keep and which one to discard, is there a way to select the edges such that we can accelerate or delay the appearance of certain substructures in the random graph? His original question concerned delaying the appearance of a so-called *giant component*, i.e. a component of size linear in the number of vertices. This was answered positively by Bohman and Frieze in [BF01].

Similarly to online Ramsey games the results one can obtain depend on the strategy employed by the player, which we call Chooser. Of particular interest is the so-called *min-product rule*, a strategy for Chooser which when comparing the two candidate edges multiplies the product of the size of the connected component incident to each endpoint of each the edge. She then selects the edge for which this product is the lowest. On the basis of computer simulations Achlioptas, D'Souza and Spencer [ADS09] conjectured that when playing with this rule the phase transition from having no giant component to one of size at least εn, $\varepsilon > 0$, is not continuous. This conjecture attracted much attention, but was finally disproved by Riordan and Warnke [RW12].

The online F-avoidance game was first studied in the Achlioptas setting by Krivelevich, Loh and Sudakov in [KLS09]. They study the game in which r edges are presented to the player and she has to immediately choose one which remains in the graph an discard the rest. Her goal is to avoid creating a copy of F for as long as possible. Krivelevich et al. determine the threshold for the appearance of cycles, cliques and complete bipartite graphs for all $r \geq 2$. They conjecture a general threshold valid for most graphs F. Mütze, Spöhel and Thomas [MST11] disproved this conjecture and proved a much more complex threshold function which is valid for arbitrary graphs F and all $r \geq 2$.

An online game at the intersection of Achlioptas processes and Ramsey-type coloring games is the *balanced Ramsey game*. As in the Achlioptas game in each round of this game the player is presented with r edges at a time, but instead of discarding all but one she has r colors at her disposal. She must immediately assign a different color to each edge, and her goal is to avoid a monochromatic copy of some fixed forbidden graph F for as long as possible. This game was introduced by Marciniszyn, Mitsche and Stojaković [MMS07], and they determine the threshold for the appearance of cycles for 2 colors. Prakash, Spöhel and Thomas [PST09] generalized these results to an arbitrary number of colors and proved thresholds valid for a large class of graphs including cycles and cliques. They also studied the vertex variant of this game and proved similar results.

Note that the F-avoidance Achlioptas game is essentially a balanced Ramsey game in which only one color is dangerous. Any winning strategy for the balanced Ramsey game can be converted to one for the Achlioptas game by simply interpreting the red edges as chosen and those in any other color as discarded. This implies that any upper bound on the threshold for the Achlioptas game also holds for the balanced game.

Due to the similarity of these two games it is a natural question to ask whether their thresholds are in fact the same or not. Krivelevich, Spöhel and Steger [KSS10] studied this for the offline setting of the Achlioptas and the balanced Ramsey game, in which the player is allowed to see all r-tuples of edges before committing to a coloring. They show that for most graphs the threshold for the two games are equal and correspond to those of the Ramsey problem in random graphs given in Theorem 3.5. For the online case they note that for F being a tree or forest there are simple examples which show that the thresholds must differ. However, they also note that all known results for non-forests show that the two thresholds are always equal. They ask whether this is true for every graph F which is not a forest.

We settle this question in Chapter 6. For the edge case we answer it negatively and show that there are graphs F which are not forests for which the two thresholds are distinct. We also consider the variant of both games in which we color vertices instead of edges. We show that in contrast to the edge case in this setting the two thresholds coincide.

Chapter 4

A randomized version of Ramsey's theorem

In this chapter we examine a version of Ramsey's theorem for random graph with a different take on randomization than Rödl and Ruciński's well known theorem. This idea was suggested in a paper by Allen, Böttcher, Hladký, and Piguet [ABHP13]. This is joint work with Yury Person, Angelika Steger and Henning Thomas. An extended abstract has appeared in [GPST11] and the full results in [GPST12].

4.1 Introduction

Recall Rödl and Ruciński's Ramsey theorem for random graphs.

Theorem 4.1 ([RR93, RR95, FK00]). *For all integers $r \geq 2$ and for every non-empty graph F which is not a forest of stars and paths of length 3 there exist constants $c > 0$ and $C > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left[G(n,p) \to (F)_r^e\right] = \begin{cases} 0 & \text{if } p \leq cn^{-1/m_2(F)}, \\ 1 & \text{if } p \geq Cn^{-1/m_2(F)}. \end{cases}$$

Note that $p = n^{-1/m_2(F)}$ is the density where we expect that every edge is contained in roughly a constant number of copies of F. This observation can be used to provide an intuitive understanding of the bounds of Theorem 4.1. If c is very small, then the number of copies of F is a.a.s. (*asymptotically almost surely*, i.e., with probability $1 - o(1)$ if n tends to infinity) small enough that they are so scattered that a coloring without a monochromatic copy of F can be found. If on the other hand C is big then these copies a.a.s. overlap so heavily that every coloring has to induce at least one monochromatic copy of F. Extending this work, Friedgut and Krivelevich [FK00] proved the thresholds of Theorem 4.1 to be sharp for most trees, and later Friedgut, Rödl, Ruciński and Tetali [FRRT06] showed that this is also the case for F being a triangle.

In this chapter we follow up on a different way to introduce randomness into Ramsey theory that was suggested in a recent paper by Allen, Böttcher, Hladký, and Piguet [ABHP13]. Note that the setup from Theorem 4.1 essentially studies the question of how many edges we need to remove from the complete graph (in a random fashion) such that the remaining graph can be colored without a monochromatic clique of size k. (For simplicity we here just consider the case $F = K_k$.) Allen et al. suggest to study the question for the case that we do not care about all cliques K_k, but only want to avoid certain cliques. More formally, assume we have a k-uniform hypergraph $H = (V_n, E(H))$ and we ask: does every coloring of the edges of the complete graph K_n on vertex set V_n with r colors induce a monochromatic k-clique *that forms a hyperedge in* H? We use the notation $K_n \xrightarrow{H} (K_k)_r^e$ to denote this property. The question asked in [ABHP13] is the following. Assume $H^k(n,q)$ is a binomial k-uniform hypergraph with edge probability q. What is the threshold $q = q(n)$ for the property $K_n \xrightarrow{H^k(n,q)} (K_k)_r^e$. In [ABHP13] the authors study the corresponding question for Turán's Theorem [Tur41]. Our first main result not only solves this problem, but – similarly to [ABHP13] – also combines it with the classical probabilistic approach by considering a random graph $G(n,p)$ instead of K_n. More precisely, we show:

Theorem 4.2 (Cliques)**.** *Let $k \geq 3$ and $r \geq 2$ be fixed integers. There exist constants $c = c(k,r) > 0$ and $C = C(k,r) > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \overset{H^k(n,q)}{\longrightarrow} (K_k)^e_r] = \begin{cases} 0, & \text{if } n^k p^{\binom{k}{2}} q \leq cn^2 p \\ 1, & \text{if } n^k p^{\binom{k}{2}} q \geq Cn^2 p \end{cases}.$$

We also study the generalization of Theorem 4.2 to strictly 2-balanced graphs F instead of cliques. A graph F is *2-balanced* if for every subgraph $J \subseteq F$ we have $d_2(J) \leq d_2(F)$, and *strictly 2-balanced* if the inequality is strict for every $J \neq F$. Note that in order to prove the 0-statement for the classical probabilistic Ramsey theorem, i.e., Theorem 4.1, we only need to consider strictly 2-balanced graphs: if we know that it holds for all strictly 2-balanced graphs then it easily follows for all other graphs F by considering an appropriate strictly 2-balanced subgraph. In the context of Theorem 4.2 such an argument is no longer trivially true. Intuitively the threshold for the $G(n,p)$ part of the theorem is determined by a strictly 2-balanced subgraph which might be distinct from F itself. The hypergraph on the other hand forbids entire copies of F and not just the 2-balanced subgraph. We thus restrict our considerations to strictly 2-balanced graphs F and leave possible further generalizations to future work.

In the following let F denote a strictly 2-balanced graph. Note that for graphs F different from cliques we will in general have more than one possible copy of F on a given vertex set of size $v(F)$. We thus need to specify which of these subgraphs are forbidden to appear monochromatically. To define this formally, we call a subset $\mathcal{R}$ of all unlabeled copies of F in K_n a *restriction set*. Then, for every graph G on n vertices we have $G \overset{\mathcal{R}}{\longrightarrow} (F)^e_r$ if every r-edge-coloring of G contains a monochromatic copy of F that is contained in $\mathcal{R}$. Moreover, let $\mathcal{R}_F(n,q)$ denote a random subset of all unlabeled copies of F in K_n in which every copy is present independently of the others with probability q.

Theorem 4.3 (Main Result)**.** *Let F be a strictly 2-balanced graph with at least 3 edges. Let $r \geq 2$ be a fixed integer. There exist constants $c = c(F,r) > 0$ and $C = C(F,r) > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \overset{\mathcal{R}_F(n,q)}{\longrightarrow} (F)^e_r)] = \begin{cases} 0, & \text{if } n^{v(F)} p^{e(F)} q \leq cn^2 p \\ 1, & \text{if } n^{v(F)} p^{e(F)} q \geq Cn^2 p \end{cases}.$$

Note that the statement is not true for strictly 2-balanced graphs with less than 3 edges. For $F = K_{1,2}$ and $q = 1$ the threshold is given by the appearance of the star $K_{1,r+1}$ in $G(n,p)$ which is known to be Poisson-distributed in the regime where p is of order $n^{1-1/(r+1)}$ (cf. [ER60]).

Moreover, note that Theorem 4.2 is a special case of Theorem 4.3. Observe also that for $q = 1$ the threshold matches with the one from Theorem 4.1, that is, our result can be viewed as a natural generalization of Theorem 4.1. The proof of the 1-statement of our results builds upon the proof of the 1-statement of Theorem 4.1. In fact, this proof applies to all graphs F that have $m_2(F) \geq 1$. Our proof of the 0-statement of Theorem 4.3 on the other hand follows a new approach by further developing algorithmic ideas from [MSSS09] to obtain a suitable coloring of $G(n,p)$ (see section 3 in [MSSS09]) and combining them with a general theorem from [RR93] about the global density of graphs which are Ramsey with respect to a given graph, i.e., when $H \nrightarrow (F)^e_2$, see Theorem 4.4.

4.2 Proof of the main result

The intuition behind the threshold in Theorem 4.3 can be stated similarly to the intuition of Theorem 4.1 as follows. Call a copy of F in a graph G *bad* with respect to a restriction set $\mathcal{R}$ if it is contained in $\mathcal{R}$. For a fixed edge in $G(n,p)$ we expect order of $n^{v(F)-2}p^{e(F)-1}q$ bad copies of F in $G(n,p)$ with respect to $\mathcal{R}_F(n,q)$ that contain this edge. Hence, if c is small and p and q satisfy the inequality of the 0-statement we expect so few bad copies of F on a fixed edge that these copies are indeed so scattered that we can find an edge-coloring without a monochromatic copy of F. However, if C is large and p and q are as in the 1-statement, then the bad copies of F overlap so heavily that no such coloring can be found.

We now prove Theorem 4.3. We first address the easier 1-statement. Throughout the remainder of this chapter, we consider F to be a fixed strictly 2-balanced graph with at least 3 edges and $r \geq 2$ to be fixed. We write $\mathcal{R}(n,q)$ instead of $\mathcal{R}_F(n,q)$.

4.2.1 The 1-statement

We have to show that for a large enough C and $n^{v(F)}p^{e(F)}q \geq Cn^2p$ the random graph $G(n,p)$ and the random restriction set $\mathcal{R}(n,q)$ a.a.s. have the property that every r-edge-coloring contains a monochromatic bad copy of F. Observe that $n^{v(F)}p^{e(F)}q \geq Cn^2p$ in particular implies $p \geq C'n^{-1/m_2(F)}$. We now use Theorem 3 from [RR95] which states that there exist constants $a, b > 0$ such that for C' large enough $p \geq C' n^{-1/m_2(F)}$ implies that every r-edge-coloring of $G(n,p)$ contains at least $an^{v(F)}p^{e(F)}$ monochromatic copies of F with probability at least $1 - 2^{-bn^2 p}$. Also note that by Chernoff bounds (see e.g. Chapter 2, Theorem 2.1 in [JŁR00]) it holds that

$$\mathbb{P}\left[e(G(n,p)) \geq 2\binom{n}{2}p\right] \leq \mathrm{e}^{-\Theta(n^2 p)}$$

and thus there are a.a.s. at most $r^{2\binom{n}{2}p}$ r-edge-colorings of $G(n,p)$. Let $\mathcal{Q}$ denote the set of all graphs for which every r-edge-coloring contains at least $an^{v(F)}p^{e(F)}$ monochromatic copies of F and which have at most $2\binom{n}{2}p$ edges. Then,

$$\mathbb{P}[G(n,p) \notin \mathcal{Q}] \leq 2^{-bn^2 p} + \mathrm{e}^{-\Theta(n^2 p)} = o(1). \tag{4.1}$$

Now, the probability that $G(n,p) \overset{\mathcal{R}(n,q)}{\not\rightarrow} (F)^e_r$ conditioned on that $G(n,p)$ satisfies $\mathcal{Q}$ can be bounded from above with a union bound over all r-edge-colorings by

$$r^{2\binom{n}{2}p}(1-q)^{an^{v(F)}p^{e(F)}} \leq \mathrm{e}^{\ln(r)n^2 p - an^{v(F)}p^{e(F)}q} \leq \mathrm{e}^{(\ln r - aC)n^2 p}, \tag{4.2}$$

where we used $n^{v(F)}p^{e(F)}q \geq Cn^2 p$ in the last step. We now have

$$\begin{aligned}
&\mathbb{P}\left[G(n,p) \overset{\mathcal{R}(n,q)}{\not\rightarrow} (F)^e_r\right] \\
&\leq \mathbb{P}\left[G(n,p) \overset{\mathcal{R}(n,q)}{\not\rightarrow} (F)^e_r \mid G(n,p) \in \mathcal{Q}\right] \cdot \mathbb{P}[G(n,p) \in \mathcal{Q}] + \mathbb{P}[G(n,p) \notin \mathcal{Q}] \\
&\overset{(4.1),(4.2)}{\leq} \mathrm{e}^{(\ln r - aC)n^2 p} + o(1).
\end{aligned}$$

Clearly, for large enough C this probability tends to 0. This finishes the proof of the 1-statement.

4.2.2 The 0-statement

Recall that we need to show that for $n^{v(F)}p^{e(F)}q \leq cn^2 p$ the random graph $G(n,p)$ and the random restriction set $\mathcal{R}(n,q)$ a.a.s. have the property that there *exists* an r-edge-coloring of $G(n,p)$ that does not contain any monochromatic bad copy of F. We call such a coloring *valid*. In the remainder we describe an algorithm that finds such a coloring and show that it succeeds with high probability.

In a first step we identify a set of edges that we can color easily. Let G be a graph and $\mathcal{R}$ be a restriction set. Let $e \in E(G)$ be an edge which is contained in at most one bad copy of F with respect to $\mathcal{R}$. Then clearly, if there exists a valid coloring for $G - e$ (the graph obtained from G by removing e), we can extend it to one for G since we can assign at least $r - 1 \geq 1$ colors to e without creating a monochromatic bad copy of F. We call such edges *open* with respect to $\mathcal{R}$ and all other edges *closed* (with respect to $\mathcal{R}$). It is easy to see that successively removing open edges yields the unique maximum subgraph of G in which every edge is contained in at least two bad copies of F, where maximum is with respect to subgraph inclusion. We call this subgraph the F*-core* of G (with respect to $\mathcal{R}$). By the above argument, it suffices to find a valid coloring for the F-core of G.

We say that a subgraph H of the F-core of G is F*-closed* with respect to $\mathcal{R}$ if every bad copy of F from the F-core of G is either contained in H or edge-disjoint with H. It is easy to see that the edges of the F-core can be partitioned into minimal F-closed subgraphs where minimal is with respect to subgraph inclusion. Furthermore, each such subgraph can be colored separately in order to find a valid coloring of the F-core. The key property of F-closed subgraphs H follows from the definition of the F-core. For every edge $e \in E(H)$ there are at least two bad copies $F_1, F_2 \subseteq H$ which contain e.

Grow Sequences

In the following we describe a procedure that yields a sequence of bad copies of F which construct an F-closed subgraph. Let F_1 be a bad copy of F from the F-core. Now, for every $\ell \geq 1$, we let $F_{\ell+1}$ be a bad copy of F from the F-core such that $F_{\ell+1}$ intersects $\bigcup_{i=1}^{\ell} F_i$ in at least one edge and for which $F_{\ell+1} \neq F_i$ for every $1 \leq i \leq \ell$. If no such copy exists in the F-core of G the sequence ends after the ℓ-th step and we set $S := (F_1, F_2, \ldots, F_\ell)$ and $G(S) := \bigcup_{i=1}^{\ell} F_i$ as *the graph of* S. We call such a sequence a *grow sequence*, and say that S is *contained* in or *appears* in G with respect to $\mathcal{R}$, meaning that $G(S)$ is a subgraph of G and that every F_i is contained in $\mathcal{R}$. Observe that for every minimal F-closed subgraph H there exists a grow sequence S such that $G(S) = H$. It thus suffices to show that we can find a valid coloring for the graph of every grow sequence that is contained in $G(n,p)$ with respect to $\mathcal{R}(n,q)$.

The following theorem by Rödl and Ruciński states that there exists a valid coloring for a graph whenever it is sparse enough. For a graph H let

$d(H) = e(H)/v(H)$ denote the density of H and let $m(H) = \max_{J \subseteq H} d(J)$ denote the maximum density of H.

Theorem 4.4 ([RR93]). *Let G and H be two graphs. If $m(H) \leq m_2(G)$ and $m_2(G) > 1$ then $H \nrightarrow (G)^e_2$.*

Note that since F is strictly 2-balanced and has at least 3 edges it satisfies $m_2(F) > 1$. Hence, we can find a valid coloring for every grow sequence S which satisfies $m(G(S)) \leq m_2(F)$. It remains to deal with grow sequences that encode denser subgraphs. The most important structural property of a grow sequence is that every edge of its graph is contained in at least two copies of F. Intuitively speaking, this forces its graph to be dense and to contain many copies of F. On the other hand, the conditions on p and q imply that $G(n,p)$ and $\mathcal{R}(n,q)$ have the property that a.a.s. every subgraph has few edges in $G(n,p)$ or few copies in $\mathcal{R}(n,q)$. In the remainder we use these two density restrictions in the following way. For the *length* of a grow sequence S we write $\ell(S)$. We show that there exists a constant L such that asymptotically almost surely

(i) there are no grow sequences in $G(n,p)$ with respect to $\mathcal{R}(n,q)$ of length more than L, and

(ii) for every grow sequence S of length at most L that appears in $G(n,p)$ with respect to $\mathcal{R}(n,q)$ we have that $G(S)$ is sparse enough so that Theorem 4.4 applies.

These two conditions are given formally in Lemma 4.5 and 4.6 below.

Lemma 4.5. *For every strictly 2-balanced graph F with at least 3 edges and every integer $r \geq 2$ there exist constants $c = c(F,r) > 0$ and $L = L(F,r) > 0$ such that if $p \leq cn^{-1/m_2(F)}q^{-1/(e(F)-1)}$ then $G(n,p)$ and $\mathcal{R}(n,q)$ a.a.s. satisfy that every grow sequence in $G(n,p)$ (with respect to $\mathcal{R}(n,q)$) has length at most L.*

Note that our assumption $n^{v(F)}p^{e(F)}q \leq cn^2p$ is equivalent to the property $p \leq c'n^{-1/m_2(F)}q^{-1/(e(F)-1)}$ for $c' = c^{1/(e(F)-1)}$.

Lemma 4.6. *For every strictly 2-balanced graph F with at least 3 edges and every integer $r \geq 2$, and for all constants $c, L > 0$ we have that if $p \leq cn^{-1/m_2(F)}q^{-1/(e(F)-1)}$ then $G(n,p)$ and $\mathcal{R}(n,q)$ a.a.s. satisfy that every grow sequence S in $G(n,p)$ (with respect to $\mathcal{R}(n,q)$) of length at most L satisfies $m(G(S)) \leq m_2(F)$.*

With these two ingredients the proof of the 0-statement from Theorem 4.3 is straightforward.

Proof (of Theorem 4.3*, 0-statement).* Choose $c = c(F, r)$ and $L = L(F, r)$ according to Lemma 4.5. Then, by lemmas 4.5 and 4.6, $G(n, p)$ and $\mathcal{R}(n, q)$ a.a.s. satisfy that every grow sequence S in $G(n, p)$ with respect to $\mathcal{R}(n, q)$ has length at most L and satisfies $m(G(S)) \leq m_2(F)$. Hence, a.a.s. every F-closed subgraph H in $G(n, p)$ with respect to $\mathcal{R}(n, q)$ satisfies $m(H) \leq m_2(F)$. Conditioning on this property of $G(n, p)$ and $\mathcal{R}(n, q)$ Theorem 4.4 guarantees that we can find a valid r-edge-coloring for every such subgraph. Moreover, the union of the valid colorings of all F-closed subgraphs yields a valid coloring of the F-core of $G(n, p)$ which can be extended to a valid coloring of $G(n, p)$. □

4.3 Proof of Lemma 4.6

Proof (of Lemma 4.6*).* As a first step of the proof we show that every grow sequence $S = (F_1, F_2, \ldots, F_\ell)$ of length at most L that satisfies $m(G(S)) > m_2(F)$ in fact satisfies $d(G(S)) > m_2(F)$. We show this by inductively constructing a finite sequence of graphs $H_0, H_1, \ldots$ all of which have density larger than $m_2(F)$ and which end in the graph $G(S)$. Let $H_0 \subseteq G(S)$ be an arbitrary densest subgraph of $G(S)$, i.e., $d(H_0) = m(G(S)) > m_2(F)$. If $H_i = G(S)$, we are done. Otherwise there exists a copy F_j in S that is neither entirely contained in H_i nor edge-disjoint from H_i. Hence, F_j overlaps with H_i in a subgraph J that contains at least one edge, that is, $e(J) \geq 1$ and $v(J) \geq 2$. We set $H_{i+1} = H_i \cup F_j$. Denote by e_i and v_i the number of new edges and new vertices, i.e.,

$$e_i = e(F) - e(J) \quad \text{and} \quad v_i = v(F) - v(J). \tag{4.3}$$

If there are no new vertices ($v_i = 0$), then clearly $d(H_{i+1}) \geq d(H_i) > m_2(F)$. Hence, we may assume from now on that there is at least one new vertex ($v_i \geq 1$). We show $e_i/v_i \geq m_2(F)$ with a case distinction.

Case 1: $e(J) = 1$. In this case,

$$\frac{e_i}{v_i} \overset{(4.3)}{=} \frac{e(F) - e(J)}{v(F) - v(J)} \overset{v(J)\geq 2}{\geq} \frac{e(F) - 1}{v(F) - 2} = m_2(F).$$

Case 2: $e(J) \geq 2$. Clearly, this implies $v(J) \geq 3$. Recall that F is strictly 2-balanced and that $J \neq F$ since there is at least one new vertex. Thus,

$e(J) - 1 < m_2(F)(v(J) - 2)$, and we can settle this case with

$$\frac{e_i}{v_i} \overset{(4.3)}{=} \frac{e(F) - e(J)}{v(F) - v(J)} = \frac{e(F) - 1 - (e(J) - 1)}{v(F) - 2 - (v(J) - 2)}$$
$$> \frac{m_2(F)(v(F) - 2) - m_2(F)(v(J) - 2)}{v(F) - 2 - (v(J) - 2)} = m_2(F). \tag{4.4}$$

Hence, we have

$$d(H_{i+1}) = \frac{e(H_{i+1})}{v(H_{i+1})} = \frac{e(H_i) + e_i}{v(H_i) + v_i} > m_2(F),$$

where in the last step we used Proposition 2.1 together with the assumption $d(H_i) = e(H_i)/v(H_i) > m_2(F)$ and $e_i/v_i \geq m_2(F)$. Clearly, after a constant number of iterations we arrive at $H_i = G(S)$.

We now show that the probability that there exists a grow sequence S of length at most L which satisfies $m(G(S)) > m_2(F)$ and which appears in $G(n,p)$ with respect to $\mathcal{R}(n,q)$ is $o(1)$. Observe that a grow sequence of length at most L can involve at most $v_L := 2 + L \cdot (v(F) - 2)$ vertices. Now fix $k \leq v_L$ and let U be a fixed vertex set of size k. We show that the probability that there is a grow sequence S of length at most L with $m(G(S)) > m_2(F)$ and $v(G(S)) = k$ that appears on U is $o(n^{-k})$. With a union bound over all k and all vertex sets U of size k this then concludes the proof of the lemma.

Observe that there are at most $\mathcal{O}(1)$ possible grow sequences that can be accommodated on the vertex set U. Consider an arbitrary but fixed possible grow sequence S of length at most L that is contained in U and that satisfies $v(G(S)) = k$ and $m(G(S)) > m_2(F)$. First observe that since $H := G(S)$ is F-closed we have that for every edge there must be two copies of F in S that contain it. Since every such copy contains $e(F)$ edges S must have length at least $2e(H)/e(F)$. Hence, using $p \leq cn^{-1/m_2(F)}q^{-1/(e(F)-1)}$ the probability that S appears on U in $G(n,p)$ with respect to $\mathcal{R}(n,q)$ can be bounded by

$$p^{e(H)}q^{2e(H)/e(F)} \leq c^{e(H)}n^{-e(H)/m_2(F)} \underbrace{q^{2e(H)/e(F) - e(H)/(e(F)-1)}}_{\leq 1, \text{ since } e(F) \geq 3}$$
$$\leq c^{e(H)}n^{-e(H)/m_2(F)} \overset{d(H) > m_2(F)}{=} o(n^{-v(H)}) = o(n^{-k}).$$

This concludes the proof of the lemma. $\square$

4.4 Proof of Lemma 4.5

Canonical Grow Sequences.

We first take a closer look at grow sequences and describe a unique canonical way to construct them. For this, let G be a graph, $\mathcal{R}$ be a restriction set, and fix an arbitrary ordering of the vertices of G. Note that this vertex ordering also naturally induces an ordering on the edges of G. Using these orderings we can say for two vertex or edge sets X and Y which of the two is *lexicographically smaller*. Now, let H be a minimal F-closed subgraph of G with respect to $\mathcal{R}$.

We can now construct a grow sequence S for H inductively in the following canonical way. Let $\mathcal{F}(H)$ denote the set of all bad copies of F in H. We set F_1 to be the lexicographically smallest copy of $\mathcal{F}(H)$. Now, for every $\ell \geq 1$, let $G(S,\ell) := \bigcup_{i=1}^{\ell} F_i$. If $G(S,\ell)$ contains an open edge with respect to the restriction set $\{F_1, F_2, \ldots, F_\ell\}$, then we let e denote the lexicographically smallest open edge. Since H is F-closed, there must be at least one bad copy of F in $\mathcal{F}(H)$ that is not in $G(S,\ell)$ and contains e. We set $F_{\ell+1}$ to be the lexicographically smallest such copy. Otherwise, if $G(S,\ell)$ only contains closed edges, then we choose $F_{\ell+1}$ to be the lexicographically smallest copy in $\mathcal{F}(H)$ which is not edge-disjoint from $G(S,\ell)$ and not yet in the sequence $F_1, F_2, \ldots, F_\ell$.

We call a grow sequence *canonical* if it follows the above procedure. Then, every F-closed subgraph has exactly one corresponding canonical grow sequence. Hence, in order to prove Lemma 4.5 it suffices to show that $G(n,p)$ a.a.s. does not contain a canonical grow sequence of size more than L with respect to $\mathcal{R}(n,q)$.

We point out here that the crucial property of canonical grow sequences is the following. In case that $G(S,i)$ has open edges we know that the copy F_{i+1} contains the lexicographically smallest of them. Roughly speaking, this property will turn out to be important in a later first moment method argument: when counting the number of grow sequences there are no choices for these two vertices.

Step Types.

Let $S = (F_1, F_2, \ldots, F_\ell)$ be a canonical grow sequence. We now view S from the perspective of building a graph step by step from $G(S,0)$ over

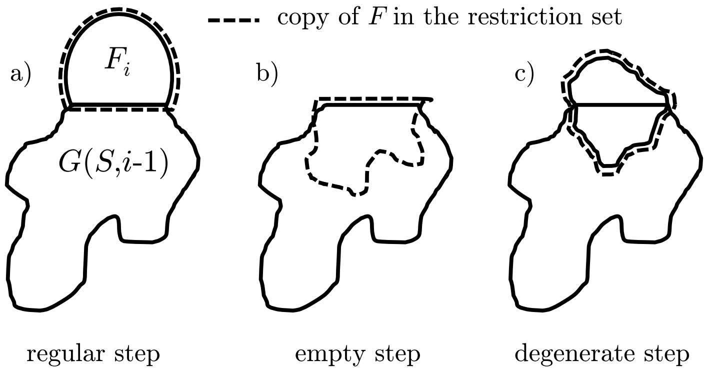


Figure 4.1: Illustration of the different step types of a grow sequence

$G(S,1)$ and so on up to $G(S)$. We split the steps of this process into different types, which are also illustrated in Figure 4.1. We call step one the *first step.* We call a step $i \geq 2$ *regular* if the intersection of F_i with $G(S,i-1)$ consists exactly of two vertices connected by an edge (Figure 4.1 *a*)), *empty* if the intersection is the whole copy F_i (Figure 4.1 *b*)) and *degenerate* otherwise (Figure 4.1 *c*)). Note that an empty step only imposes a copy of F in the restriction set. Besides categorizing the steps into regular, empty and degenerate ones we introduce another categorization. We call step i *open* if $G(S,i-1)$ contains open edges with respect to the restriction set $\{F_1, F_2, \ldots, F_{i-1}\}$ and *closed* otherwise. Note that since S is canonical the copy F_i of an open step i contains the lexicographically smallest open edge in $G(S,i-1)$.

Let v_i denote the number of new vertices added in step i, i.e., $v_i := v(G(S,i)) - v(G(S,i-1))$ and let similarly e_i denote the number of new edges. It is easy to see that for every regular and every empty step we have $v_i - e_i/m_2(F) = 0$, and that for every degenerate step $v_i - e_i/m_2(F) < 0$ since F is strictly 2-balanced (using a calculation similar to (4.4)). Furthermore, since the number of possible intersections of F_i with $G(S,i-1)$ is constant, there exists a constant $\delta = \delta(F) > 0$ such that every degenerate step satisfies

$$v_i - e_i/m_2(F) < -\delta. \tag{4.5}$$

Before we continue we give an intuitive reasoning of how we will use (4.5). We will use a first moment method to show that a.a.s. grow sequences of length more than L do not appear. In order to do so we need to *count* grow sequences. To do that we will distinguish grow sequences according to their number of degenerate steps. Observe that for a degenerate step we essentially have to choose v_i new vertices (at most n^{v_i} ways) and require the presence of e_i new edges and a new copy of F in the restriction set (which happens with probability $p^{e_i}q$). As $p \leq cn^{-1/m_2(F)}q^{-1/(e(F)-1)}$, we see that (4.5) implies that the probability for K degenerate steps decreases with $n^{-\delta K}$. That is, we should not expect that we have grow sequences with 'many' degenerate steps. In order to handle grow sequences with 'few' degenerate steps we will now prove some deterministic properties of canonical grow sequences.

We show that long sequences have to contain a certain number of degenerate steps, and that the number of closed regular steps and empty steps can be bounded in terms of the number of degenerate steps. For a grow sequence S we use $\mathrm{degen}(S), \mathrm{empty}(S), \mathrm{regular}^o(S), \mathrm{regular}^c(S)$ to denote the number of degenerate, empty, open regular and closed regular steps in S. Note that $\mathrm{degen}(S)+\mathrm{empty}(S)+\mathrm{regular}^o(S)+\mathrm{regular}^c(S) = \ell(S)-1$, where the -1 accounts for the first step. Moreover, we write $X_F(G)$ for the number of (not necessarily bad) copies of F in G.

Claim 4.7. *Let $S = (F_1, F_2, \ldots, F_\ell)$ be a grow sequence. If step i is empty or regular, then $X_F(G(S,i)) \leq X_F(G(S,i-1)) + 1$.*

Proof. The claim is trivial if step i is an empty step. Otherwise step i is regular, and we can denote by $e = \{u, v\}$ the intersection of F_i with $G(S, i-1)$. Assume that in addition to the copy F_i there is another copy $\tilde{F}$ of F created in that step. Then clearly, $\tilde{F}$ contains at least one edge from the new edges of step i, i.e., from $E(F_i) \setminus \{e\}$, and at least one edge from the old edges $E(G(S,i-1)) \setminus \{e\}$. Hence, the graphs $\tilde{F}_{\mathrm{new}} = \tilde{F}[V(F_i)]$ and $\tilde{F}_{\mathrm{old}} = \tilde{F}[V(G(S,i-1))]$ are both non-empty. Since every strictly 2-balanced graph is 2-vertex-connected it follows that both u and v are contained in $V(\tilde{F})$, $V(\tilde{F}_{\mathrm{new}})$ and $V(\tilde{F}_{\mathrm{old}})$. Moreover, removing u and v disconnects $\tilde{F}$.

Case 1: $e \in E(\tilde{F})$. Then we have

$$m_2(F) = m_2(\tilde{F}) = \frac{e(\tilde{F})-1}{v(\tilde{F})-2} = \frac{e(\tilde{F}_{\mathrm{new}})-1+e(\tilde{F}_{\mathrm{old}})-1}{v(\tilde{F}_{\mathrm{new}})-2+v(\tilde{F}_{\mathrm{old}})-2} < m_2(F),$$

where the last step follows from Proposition 2.1 and the fact that F is strictly 2-balanced. Clearly, this is a contradiction.

Case 2: $e \notin E(\tilde{F})$. Then we have

$$m_2(F) = m_2(\tilde{F}) = \frac{e(\tilde{F}) - 1}{v(\tilde{F}) - 2} = \frac{e(\tilde{F}_{\text{new}}) + e(\tilde{F}_{\text{old}}) - 1}{v(\tilde{F}_{\text{new}}) - 2 + v(\tilde{F}_{\text{old}}) - 2}. \tag{4.6}$$

Since F is strictly 2-balanced we have $(e(\tilde{F}_{\text{old}})-1)/(v(\tilde{F}_{\text{old}})-2) < m_2(F)$. Hence, (4.6) together with Proposition 2.1 implies

$$\frac{e(\tilde{F}_{\text{new}}) + 1 - 1}{v(\tilde{F}_{\text{new}}) - 2} = \frac{e(\tilde{F}_{\text{new}})}{v(\tilde{F}_{\text{new}}) - 2} > m_2(F).$$

Since $\tilde{F}_{\text{new}} \cup \{e\}$ is a subgraph of F_i with $e(\tilde{F}_{\text{new}}) + 1$ edges and $v(\tilde{F}_{\text{new}})$ vertices, this contradicts the property that F is strictly 2-balanced. □

Claim 4.8. *For every canonical grow sequence $S = (F_1, \dots, F_\ell)$ it holds that* $\text{regular}^c(S) \le \text{degen}(S)$. *Moreover, for every prefix sequence $S_i = (F_1, \dots, F_i)$, where $i \le \ell$, we have* $\text{regular}^c(S_i) \le \text{degen}(S_i)$.

Proof. Let $S = (F_1, F_2, \dots, F_\ell)$ be a canonical grow sequence, and let $r_1, r_2, \dots$ denote the indices of its closed regular steps and set $r_0 = 1$ for convenience. We show that for every $i \ge 0$ there is at least one degenerate step between steps r_i and r_{i+1}. Clearly, $G(S, r_i)$ contains $e(F) - 1$ open edges with respect to the restriction set $\{F_1, F_2, \dots, F_{r_i}\}$ and thus, step $r_i + 1$ is open. Moreover, since r_i is a closed step all edges in $G(S, r_i - 1)$ are closed which together with Claim 4.7 implies that the *only* copy of F in $G(S, r_i)$ that contains open edges is F_{r_i}. Since S is canonical, step $r_i + 1$ cannot be empty (an empty step would require a copy of F in $G(S, r_i)$ that contains open edges and is not yet in S). Step $r_i + 1$ is thus either degenerate or an open regular step. In the former case we are done, and in the latter case we can apply the above argument to obtain that step $r_i + 2$ is either degenerate or open regular and so on. Eventually, there must be a degenerate step between steps r_i and r_{i+1} since $G(S, r_{i+1} - 1)$ does not contain open edges with respect to $\{F_1, F_2, \dots, F_{r_{i+1}-1}\}$. □

Claim 4.9. *There exist constants $c_1 = c_1(F)$ and $c_2 = c_2(F)$ such that for every canonical grow sequence S we have*

$$\text{empty}(S) \le \big(c_1 \cdot \text{degen}(S)\big)^{c_2}.$$

Proof. Let $S = (F_1, F_2, \dots, F_\ell)$ be a canonical grow sequence. Assume that the i-th step of the sequence S is an empty step. Then $F_i \subseteq G(S, i-1)$

and $F_i \neq F_j$ for every $1 \leq j \leq i-1$. Note that such a copy can only be created in a degenerate step since by Claim 4.7 the only copy of F created in a regular step is the copy of the step itself. Hence, F_i must contain an edge that was added in a degenerate step. We call such edges *degenerate*. Let $X_{F,e}(G)$ denote the number of copies of F in G that contain edge e. Then we have

$$\text{empty}(S) \leq \sum_{\substack{e \in E(G(S)) \\ e \text{ is degenerate}}} X_{F,e}(G(S)).$$

Since there are at most $e(F)\text{degen}(S)$ degenerate edges in $G(S)$ it clearly suffices to bound $X_{F,e}(G(S))$ for every edge e by a bound similar to the one in the claim statement. Instead of bounding this quantity directly, let C_k^e denote for every edge $e \in E(G(S))$ the number of *induced* cycles of length k in $G(S)$ (where $3 \leq k \leq v(F)$) that contain e. Moreover, let C_k denote the maximum over all edges in $G(S)$. We will show that there exist constants $c_1' = c_1'(F)$ and $c_2' = c_2'(F)$ such that

$$C_k \leq (c_1' \cdot \text{degen}(S))^{c_2'}. \tag{4.7}$$

If we assume (4.7), then we can bound $X_{F,e}(G(S))$ for a fixed edge $e \in E(G(S))$ as follows. Let $\mathcal{F}$ denote the set of all supergraphs of F on $v(F)$ vertices. Clearly, for every copy of F in $G(S)$ that contains e, there is also an *induced* copy of some $J \in \mathcal{F}$ in $G(S)$ that contains e. Moreover, such an induced copy J can accommodate only a constant number of copies of F that contain e. Let $\bar{X}_{J,e}(G)$ denote the number of induced copies of J in G that contain edge e. As $|\mathcal{F}|$ is bounded by a constant (depending on F) it suffices to bound $\bar{X}_{J,e}(G)$ for every $J \in \mathcal{F}$ with a bound similar to the one in the claim statement.

Let $J \in \mathcal{F}$ be fixed. For every edge $e' \in E(J)$ fix an induced cycle $C(e')$ in J that contains e'. Note that this is possible since F is 2-connected. Clearly, the union of all these $e(J) \leq v(F)^2$ cycles covers $E(J)$, that is,

$$\bigcup_{e' \in E(J)} E(C(e')) = E(J).$$

We can now count the number of induced copies of J in $G(S)$ that contain e as follows. First choose the role of e in the copy of J ($e(J)$ possibilites). Say e has the role of $e' \in E(J)$. Then we have to choose an induced copy of $C(e')$ in $G(S)$ that contains e (at most $(c_1' \cdot \text{degen}(S))^{c_2'}$ possibilities by (4.7)). Continue with some other edge of that cycle and choose its role in the copy of J, and then choose a copy of the corresponding induced cycle on it and so on. In this way we obtain

$$\begin{aligned}\bar{X}_{J,e}(G(S)) &\leq \left(e(J)\cdot(c_1'\cdot\mathrm{degen}(S))^{c_2'}\right)^{e(J)}\\ &\overset{e(J)\leq v(F)^2}{\leq} (c_1''\cdot\mathrm{degen}(S))^{c_2''}\end{aligned}$$

for appropriately chosen constants $c_1''=c_1''(F)$ and $c_2''=c_2''(F)$. As explained above this concludes the proof.

It remains to show (4.7). For every $0\leq i\leq \ell$ let $S_i=(F_1,F_2,\ldots,F_i)$ denote the prefix of S including the first i steps. For every k let $P^{k,i}_{\{u,v\}}$ denote the number of induced paths of length k with endpoints u and v in $G(S_i)-uv$ for $k\geq 2$ and in $G(S_i)$ for $k=1$. (A path of length k is a path with k edges.) Denote with $P^{k,i}$ the maximum of $P^{k,i}_{\{u,v\}}$ over all vertex pairs $\{u,v\}$. We will show that there exists a constant $\beta=\beta(F)>0$ such that for every $1\leq k\leq v(F)-1$ and every $1\leq i\leq\ell$ we have

$$P^{k,i}\leq(\beta(\mathrm{degen}(S_i)+1))^{k-1}. \tag{4.8}$$

As $C_k\leq P^{k-1,\ell}$ for all $3\leq k\leq v(F)$ this establishes (4.7). It thus remains to show (4.8). First observe that (4.8) is trivially true for $k=1$ as $P^{1,i}\leq 1$ for every i and we therefore assume $k\geq 2$ in the remainder. Let $\{u,v\}$ be a fixed vertex pair. We will show that if β is large enough, then we have for every $k\geq 2$ and every $1\leq i\leq\ell$ that

$$P^{k,i}_{\{u,v\}}\leq\frac{\beta}{2}\left(\mathrm{regular}^{\{u,v\}}(S_i)+1\right)+6e(F)\sum_{j=1}^{\mathrm{degen}(S_i)+1}(\beta j)^{k-2}, \tag{4.9}$$

where $\mathrm{regular}^{\{u,v\}}(S_i)$ denotes the number of regular steps $j\leq i$ in which we attach a copy of F to the edge $\{u,v\}$, that is, for which $V(F_j)\cap V(G(S,j-1))=\{u,v\}$. First we show that (4.9) implies (4.8). For this, observe that $V(F_j)\cap V(G(S,j-1))=\{u,v\}$ is satisfied for at most one open regular step and at most $\mathrm{regular}^c(S_i)$ closed regular steps $j\leq i$, and hence by Claim 4.8, $\mathrm{regular}^{\{u,v\}}(S_i)\leq\mathrm{degen}(S_i)+1$. Using $k\leq v(F)$ we thus easily conclude that for a sufficiently large β we have that (4.9) implies (4.8). It thus remains to show (4.9) for every $k\geq 2$ and every $1\leq i\leq\ell$. For this, we do an induction on k and i. More precisely, we will use that for all smaller values of k we have (4.8) (the trivial case $k=1$

serves as induction base) and for all smaller values of i we have (4.9). We make a case distinction according to the step type of step i.

Case 1. Step i is regular and $P^{k,i-1}_{\{u,v\}} = 0$, or $i = 1$. Using that $P^{k,i}_{\{u,v\}}$ counts *induced* paths, it is not hard to see that $P^{k,i}_{\{u,v\}}$ can only be non-zero if at least one of the two vertices $\{u, v\}$ lies within the new vertices added in step i, that is, if $\{u,v\}\cap(V(F_i)\setminus V(G(S,i-1))) \neq \emptyset$. Clearly, if both u and v lie within $V(F_i)$, there can be at most $v(F)^k \leq v(F)^{v(F)}$ induced paths of length k between u and v and thus (4.9) is satisfied if $\beta \geq 2v(F)^{v(F)}$. Otherwise u must lie within the old vertices, i.e., $u \in V(G(S,i-1))\setminus V(F_i)$ and v in the new vertices, or vice versa. Then, we can count the number of induced paths of length k from u to v as follows (see also Figure 4.2 a)). Let $\{x, y\} = V(F_i) \cap V(G(S, i-1))$, and observe that in this case $\{u, v\} \cap \{x, y\} = \emptyset$. Now, every such path consists of an induced path of length s from u to x or y for some $1 \leq s \leq k-1$ and an induced path of length $k-s$ from x or y to v (clearly, there are at most $v(F)^{v(F)}$ such paths since they are contained in F_i). Note that paths containing both x and y can also be expressed this way by simply including the edge $\{x, y\}$ in one of the two subpaths of length s and $k-s$. Hence, we obtain

$$\begin{aligned}
P^{k,i}_{\{u,v\}} &\leq \sum_{s=1}^{k-1} P^{s,i}_{\{u,x\}} P^{k-s,i}_{\{x,v\}} + \sum_{s=1}^{k-1} P^{s,i}_{\{u,y\}} P^{k-s,i}_{\{y,v\}} \\
&\leq 4P^{k-1,i} + 2v(F)^{v(F)} \sum_{s=2}^{k-2} P^{s,i} \\
&\overset{(4.8),k\leq v(F)}{\leq} 4(\beta(\mathrm{degen}(S_i)+1))^{k-2} + 2v(F)^{v(F)+1}(\beta(\mathrm{degen}(S_i)+1))^{k-3} \\
&\leq 6(\beta(\mathrm{degen}(S_i)+1))^{k-2},
\end{aligned}$$

where the last step holds whenever β is sufficiently large. Clearly, this implies (4.9).

Case 2. Step i is regular and $P^{k,i-1}_{\{u,v\}} \neq 0$. It is easy to see that in this case we can only have $P^{k,i}_{\{u,v\}} > P^{k,i-1}_{\{u,v\}}$ if the copy F_i is attached to the edge $\{u, v\}$, that is, $V(F_i) \cap V(G(S, i-1)) = \{u, v\}$. Hence, $\mathrm{regular}^{\{u,v\}}(S_i) = \mathrm{regular}^{\{u,v\}}(S_{i-1})+1$ and step i can create at most $v(F)^k \leq v(F)^{v(F)}$ new induced paths of length k between u and v. Thus (4.9) certainly remains true if $\beta \geq 2v(F)^{v(F)}$.

Case 3. Step i is degenerate. In this case we insert the edges of F_i one by one. So assume a single edge $\{x, y\}$ is inserted. Clearly, this can only create

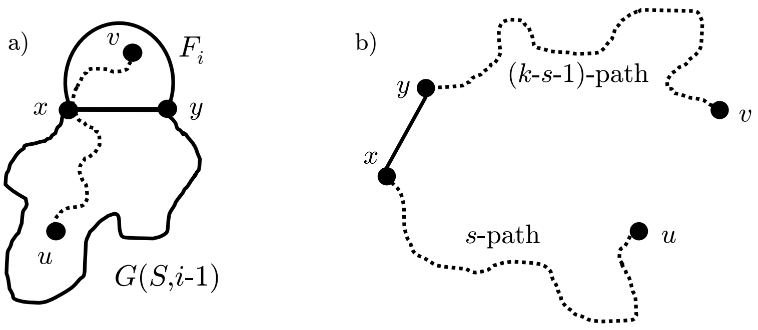


Figure 4.2: Creating a new induced path of length k from u to v

a new induced path of length k from u to v if for some $0 \leq s \leq k-1$ we have that there is an induced path of length s from u to x and an induced path of length $k-s-1$ from v to y or similarly for paths from u to y and from v to x (see Figure 4.2 b)). Observe that the case $\{x,y\} = \{u,v\}$ does not create any new induced paths of length k since $k \geq 2$. Furthermore, the case $|\{x,y\} \cap \{u,v\}| = 1$ corresponds to the case $s = 0$ or $s = k-1$, and we set $P^{0,j} = 1$ for convenience. Then, the number of induced paths of length k from u to v that were created by the insertion of $\{x,y\}$ is at most

$$\begin{aligned}\sum_{s=0}^{k-1} P^{s,j}_{\{u,x\}} P^{k-s-1,j}_{\{v,y\}} + \sum_{s=0}^{k-1} P^{s,j}_{\{u,y\}} P^{k-s-1,j}_{\{v,x\}} \\ &\leq 4P^{k-1,j} + 2\sum_{s=1}^{k-2} P^{s,j} P^{k-s-1,j} \\ &\leq 6(\beta(\text{degen}(S_i)+1))^{k-2},\end{aligned}$$

cf. the calculation above. Since we insert at most $e(F)$ edges in step i we thus have

$$P^{k,i}_{\{u,v\}} \leq P^{k,i-1}_{\{u,v\}} + e(F)6(\beta(\text{degen}(S_i)+1))^{k-2}.$$

As the assumption of this case is that step i is degenerate this implies that (4.9) holds in this case as well. This thus concludes the proof of the claim $\square$

Claim 4.10. *For every $d \geq 1$ there exists a constant $\ell_{\max}(d)$ such that every canonical grow sequence with at most d degenerate steps has length at most $\ell_{\max}(d)$.*

Proof. Let $d \geq 1$ be a constant and let S be a grow sequence of length ℓ. By Claim 4.9, there can be at most $C' := (c_1 d)^{c_2} + d + 1$ steps that are not regular. Hence, S contains at least $\ell - C'$ regular steps each of which increases the number of open edges by at least $e(F) - 2$. Moreover, every non-regular step decreases the number of open edges by at most $e(F)$. Hence, since $G(S)$ is F-closed we have $(\ell - C')(e(F) - 2) \leq C' e(F)$ and thus $\ell \leq \frac{e(F)}{e(F)-2} C' + C'$. □

We can now turn to the proof of Lemma 4.5.

Proof (of Lemma 4.5). Let $\mathcal{S}$ be the set of all canonical grow sequences of length more than $L = L(F)$ (we will fix this constant later). Note that by a first moment argument it suffices to show that for an appropriate constant $c = c(F)$ we have that if $p \leq cn^{-1/m_2(F)} q^{-1/(e(F)-1)}$ the expected number of sequences from $\mathcal{S}$ contained in $G(n, p)$ with respect to $\mathcal{R}(n, q)$ is $o(1)$.

Note that if we can show for a sequence $S \in \mathcal{S}$ that a *prefix sequence* of it, i.e., a sequence obtained by considering only the first k steps for some $1 \leq k \leq \ell(S)$, is not contained in $G(n, p)$ with respect to $\mathcal{R}(n, q)$ then S as well does not appear. Hence, if we can find a set $\mathrm{Pre}(\mathcal{S})$ such that all sequences of $\mathcal{S}$ have a prefix sequence in $\mathrm{Pre}(\mathcal{S})$ and such that a.a.s. no sequence from $\mathrm{Pre}(\mathcal{S})$ appears in $G(n, p)$ with respect to $\mathcal{R}(n, q)$ then we are done.

With this in mind we define $\mathrm{Pre}(\mathcal{S})$ as the set of the following prefixes. Let $d_{\max} = d_{\max}(F)$ be a constant which we will determine later on. For each $S \in \mathcal{S}$ we include the prefix sequence of S containing either all steps up to (and including) the $d_{\max}$-th degenerate step or all steps up to the $\log n$-th step, if the index of the $d_{\max}$-th degenerate step is larger than $\log n$. Note that this is well defined as we can force any sequence $S \in \mathcal{S}$ to contain at least $d_{\max}$ many degenerate steps by choosing L large enough, cf. Claim 4.10.

The key intuition is that prefixes $S \in \mathrm{Pre}(\mathcal{S})$ containing 'many' degenerate steps give rise to a very dense graph $G(S)$, which correspondingly is unlikely to appear. On the other hand prefixes $S \in \mathrm{Pre}(\mathcal{S})$ with 'few' degenerate steps must contain many regular steps. The corresponding graph $G(S)$ will then be very large and also unlikely to appear.

As each prefix sequence in $\mathrm{Pre}(\mathcal{S})$ contains at most $d_{\max}$ degenerate steps we have by Claim 4.8 and Claim 4.9 that they also contain at most $d_{\max}$ closed regular and at most $(c_1 d_{\max})^{c_2} =: e_{\max}$ empty steps. Let $m = 2d_{\max} + e_{\max}$ denote the maximum number of steps that are not open regular in any prefix sequence. Note that m is a fixed constant depending only on F. Hence, we can choose L such that $L > m$ holds.

We define $\mathrm{Pre}^{d_{\max}}(\mathcal{S})$ as the set containing all prefix sequences from $\mathrm{Pre}(\mathcal{S})$ with exactly $d_{\max}$ degenerate steps and $\mathrm{Pre}^{\log n}(\mathcal{S})$ as the set of those with length exactly $\log n$ and less than $d_{\max}$ many degenerate steps. Clearly, every prefix sequence in $\mathrm{Pre}(\mathcal{S})$ is in at least one of the two subsets. We consider both subsets separately.

Sequences with $\log n$ steps

We start with prefix sequences in $\mathrm{Pre}^{\log n}(\mathcal{S})$. We can bound the number of elements in $\mathrm{Pre}^{\log n}(\mathcal{S})$ by counting all sequences of steps of length $\log n$ which contain at most m steps that are not open regular. To do so we first fix the number of steps that are not open regular, their types (i.e., open or closed, and regular, empty, degenerate) and their position in the sequence. For this we have at most

$$m(5\log n)^m \tag{4.10}$$

choices. Then for any sequence with a fixed configuration of steps we have at most $n^{v(F)}$ choices for the copy of a step that is not open regular. As there are at most m of these we have in total at most

$$n^{v(F)m} \tag{4.11}$$

different choices for these steps. All remaining steps are open regular steps. Recall that for every open step two vertices of its copy of F are determined by all previous steps since all sequences in $\mathcal{S}$ are canonical. So for each of the at most $\log n$ open regular steps we only need to choose $v(F) - 2$ new vertices and the role of the two predetermined vertices in the copy of F. Thus, every open regular step gives at most

$$v(F)^2 n^{v(F)-2} \tag{4.12}$$

choices. The very first step is special and we model it by choosing two vertices as the starting edge (n^2 choices), and another $v(F)^2 n^{v(F)-2}$ choices as in an open regular step. Together with (4.10), (4.11) and (4.12) we therefore have that the number of elements in $\mathrm{Pre}^{\log n}(\mathcal{S})$ is at most

$$m\big(5\log(n)\big)^m \cdot n^{v(F)m} \cdot n^2 \cdot \big(v(F)^2 n^{v(F)-2}\big)^{\log n}.$$

For a fixed sequence in $\mathrm{Pre}^{\log n}(S)$ the probability that it appears in $G(n,p)$ with respect to $\mathcal{R}(n,q)$ is bounded from above by the probability that the first $\log n - m$ open regular steps appear. Each such step requires $e(F)-1$ new edges to be present in $G(n,p)$ and one new copy of F to be present in $\mathcal{R}(n,q)$, so this probability is at most $(p^{e(F)-1}q)^{\log n - m}$. Using $n^{v(F)-2}p^{e(F)-1}q \le c$ we can now deduce that the number $X_{\mathrm{Pre}^{\log n}(\mathcal{S})}$ of sequences from $\mathrm{Pre}^{\log n}(\mathcal{S})$ that appear in $G(n,p)$ with respect to $\mathcal{R}(n,q)$ satisfies

$$\begin{aligned}
\mathbb{E}[X_{\mathrm{Pre}^{\log n}(\mathcal{S})}] &\le m\big(5\log(n)\big)^m \cdot n^{v(F)m} \cdot n^2 \\
&\quad \cdot \big(v(F)^2 n^{v(F)-2}\big)^{\log n} \cdot (p^{e(F)-1}q)^{\log n - m} \\
&\le m\big(5\log(n)\big)^m \cdot n^{2v(F)m+2} \cdot v(F)^{2\log n} \cdot c^{\log n - m} \\
&= o(n) \cdot n^{2v(F)m+2+2\log(v(F))-\log(1/c)}.
\end{aligned}$$

As m is a constant depending only on F we can choose $c = c(F)$ such that the above expectation is $o(1)$.

Sequences with $d_{\max}$ degenerate steps

It remains to consider the prefix sequences in $\mathrm{Pre}^{d_{\max}}(\mathcal{S})$, i.e., those which contain exactly $d_{\max}$ degenerate steps. We partition these further into sets $\mathrm{Pre}^{d_{\max}}(V,E,O,\ell)$ which contain all sequences from $\mathrm{Pre}^{d_{\max}}(\mathcal{S})$ of length exactly ℓ for which the total number of new vertices and new edges added in the $d_{\max}$ degenerate steps are exactly V and E respectively, and for which the number of empty steps is exactly O. Clearly, this set can only be non-empty if $1 \le V \le d_{\max}(v(F)-3)$, $1 \le E \le d_{\max}(e(F)-2)$, $0 \le O \le e_{\max}$ and $1 \le \ell \le \log n$. Hence, the total number of subsets that we need to consider is bounded by

$$d_{\max}^2(v(F)-3)(e(F)-2)e_{\max}\log n \le \log(n)^2 \tag{4.13}$$

for n large enough.

Recall that for every degenerate step the numbers v_{new} of new vertices and e_{new} of new edges satisfy $v_{\mathrm{new}} - e_{\mathrm{new}}/m_2(F) < -\delta$ where $\delta = \delta(F) > 0$, cf. (4.5). Therefore the set $\mathrm{Pre}^{d_{\max}}(V, E, O, \ell)$ can only be non-empty if

$$V - E/m_2(F) < -\delta d_{\max}. \tag{4.14}$$

We now derive a bound on the number of sequences contained in a set $\mathrm{Pre}^{d_{\max}}(V, E, O, \ell)$. Similar to (4.10) we have at most

$$m(5\ell)^m \tag{4.15}$$

choices for the step configuration, i.e., the number of steps that are not open regular, their types and positions. Moreover, we again model the first step by choosing a starting edge (n^2 choices) and by counting the remaining choices similar to an open regular step.

For each of the O empty steps in a sequence we need to choose a copy of F within the *old* vertices, i.e., the vertices which have appeared previously in the sequence. As the entire sequence contains at most $v(F)\ell$ vertices we have at most

$$(v(F)\ell)^{v(F)O}$$

choices in total.

Considering all degenerate steps at once we need to choose a total of $d_{\max} v(F)$ vertices for them, V of which are new vertices and not from the previously seen ones, resulting in at most n^V choices, and $d_{\max} v(F) - V \leq d_{\max} v(F)$ of which are chosen from the previously seen ones, giving at most $(v(F)\ell)^{d_{\max} v(F)}$ choices. Having fixed the new vertices and old vertices it remains to choose for every degenerate step which vertices of the copy of F are from the old and which from the new vertices. This gives at most another $2^{v(F)d_{\max}} \leq (v(F)\ell)^{v(F)d_{\max}}$ choices. In total the number of choices for degenerate steps is bounded by

$$n^V (v(F)\ell)^{2v(F)d_{\max}}.$$

All other $\ell - d_{\max} - O$ steps are regular steps. Similar to (4.12) each open regular step gives rise to at most $\left(v(F)^2 n^{v(F)-2}\right)$ choices. For the at most $d_{\max}$ closed regular steps we additionally have to select the two vertices that in contrast to open regular steps are not predetermined. This accounts at most for another $(v(F)\ell)^{2d_{\max}}$ choices. In total all regular steps give rise to at most

$$(v(F)\ell)^{2d_{\max}} \left(v(F)^2 n^{v(F)-2}\right)^{\ell - d_{\max} - O} \tag{4.16}$$

choices. Combining (4.15)–(4.16) and the n^2 choices for the starting edge, we get that the number of sequences in the set $\mathrm{Pre}^{d_{\max}}(V,E,O,\ell)$ is at most

$$\begin{aligned} m(5\ell)^m &\cdot n^2 \cdot (v(F)\ell)^{v(F)O} \cdot n^V (v(F)\ell)^{2v(F)d_{\max}} \\ &\cdot (v(F)\ell)^{2d_{\max}} \left(v(F)^2 n^{v(F)-2}\right)^{\ell-d_{\max}-O} \\ \le\; & \underbrace{m(5v(F)\ell)^{m+v(F)O+2v(F)d_{\max}+2d_{\max}}}_{=o(n)} \cdot \underbrace{v(F)^{2\ell}}_{\le n^{2\log(v(F))}} \\ &\cdot n^{2+V+(v(F)-2)(\ell-d_{\max}-O)} \\ \le\; & n^{3+2\log(v(F))+V+(v(F)-2)(\ell-d_{\max}-O)}, \end{aligned} \tag{4.17}$$

where the last step holds for n large enough. It is easy to see that a prefix sequence from the set $\mathrm{Pre}^{d_{\max}}(V,E,O,\ell)$ appears in $G(n,p)$ with respect to $\mathcal{R}(n,q)$ with probability

$$q^{\ell}\left(p^{e(F)-1}\right)^{\ell-d_{\max}-O} p^E. \tag{4.18}$$

Combining (4.17) and (4.18) with the fact that $n^{v(F)-2}p^{e(F)-1}q \le c$ and $p \le n^{-1/m_2(F)}q^{-1/(e(F)-1)}$, which holds if we choose $c \le 1$, we obtain that for every subset $\mathrm{Pre}^{d_{\max}}(V,E,O,\ell)$ the number $X_{\mathrm{Pre}^{d_{\max}}(V,E,O,\ell)}$ of sequences that are present in $G(n,p)$ with respect to $\mathcal{R}(n,q)$ satisfies

$$\begin{aligned} \mathbb{E}[X_{\mathrm{Pre}^{d_{\max}}(V,E,O,\ell)}] &\le n^{3+2\log(v(F))+V+(v(F)-2)(\ell-d_{\max}-O)} \\ &\quad \cdot q^{\ell}\left(p^{e(F)-1}\right)^{\ell-d_{\max}-O} p^E \\ &\le n^{3+2\log(v(F))+V} p^E q^{d_{\max}+O} \underbrace{c^{\ell-d_{\max}-O}}_{\le 1} \\ &\le n^{3+2\log(v(F))+V-E/m_2(F)} \underbrace{q^{-(E/(e(F)-1))+d_{\max}+O}}_{\le 1 \text{ since } E \le d_{\max}(e(F)-2)} \\ &\overset{(4.14)}{\le} n^{3+2\log(v(F))-\delta d_{\max}}. \end{aligned}$$

Since this bound is independent of V, E, O and ℓ, we obtain with (4.13) that we have for large enough n that

$$\mathbb{E}[X_{\mathrm{Pre}^{d_{\max}}(\mathcal{S})}] \le \log(n)^2 n^{3+2\log(v(F))-\delta d_{\max}} \le n^{4+2\log(v(F))-\delta d_{\max}}. \tag{4.19}$$

Choosing $d_{\max} = d_{\max}(F)$ large enough and $L = L(F)$ such that every canonical grow sequence with at least L steps has at least $d_{\max}$ degenerate steps (cf. Claim 4.10) we have that the expectation in (4.19) is $o(1)$. $\square$

Chapter 5

Lower bounds for Ramsey properties in random hypergraphs

In this chapter we prove a lower bound for the Ramsey problem in random hypergraphs. For a large class of hypergraphs this lower bound matches the upper bound proved by Friedgut, Rödl and Schacht [FRS10] and independently by Conlon and Gowers [CG10]. This is joint work with Yury Person, Angelika Steger and Henning Thomas, and is currently unpublished.

5.1 Introduction

Recall that Rödl and Ruciński [RR93, RR94, RR95] determined for an arbitrary graph F the threshold for the property $G(n,p) \to (F)^e_r$. To state their full results we need the following definition, given here in its more general form for k-uniform hypergraphs. Set

$$d_k(G) := \begin{cases} 0 & \text{if } e(G) = 0 \\ 1/k & \text{if } e(G) = 1, v(G) = k \\ \frac{e(G)-1}{v(G)-k} & \text{otherwise,} \end{cases} \quad \text{and} \quad m_k(G) := \max_{H \subseteq G} d_k(H). \tag{5.1}$$

We call $m_k(G)$ the *k-density* of G. If G maximizes (5.1), we call it *k-balanced*, and *strictly k-balanced* if all strict subgraphs of G have strictly lower k-density.

Theorem 5.1 ([RR93, RR95, FK00])**.** *Let F be a graph with at least one edge and $r \geq 2$.*

i) If F is a forest of stars, then

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \to (F)^e_r] = \begin{cases} 0 & \text{if } p \ll n^{-1-1/(r(\Delta(F)-1)+1)} \\ 1 & \text{if } p \gg n^{-1-1/(r(\Delta(F)-1)+1)} \end{cases}$$

ii) If $r = 2$ and F is a forest of stars and at least one path on 3 edges, then there exists a constant C such that

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \to (F)^e_r] = \begin{cases} 0 & \text{if } p \ll n^{-1/m_2(F)} = n^{-1} \\ 1 & \text{if } p \geq Cn^{-1/m_2(F)} = Cn^{-1} \end{cases} \tag{5.2}$$

iii) In all other cases there exist constants $c = c(F, r)$ and $C = C(F, r)$ such that

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \to (F)^e_r] = \begin{cases} 0 & \text{if } p \leq cn^{-1/m_2(F)} \\ 1 & \text{if } p \geq Cn^{-1/m_2(F)} \end{cases} \tag{5.3}$$

Note that for a “nice” graph F the expected number of copies of F in $G(n,p)$ is $\Theta(n^{v(F)}p^{e(F)})$, while the expected number of edges in $G(n,p)$ is $\Theta(n^2p)$. This result essentially states that the transition from the 0 to the 1 statement happens for values of p such that these two quantities are roughly equal. In other words if the expected number of copies of F per edges is smaller than some small constant c', then coloring without monochromatic F is possible, while if this number if larger than a large constant C' then a monochromatic F always appears. This matches the intuition that if there are very few copies of F on each edge, then these copies cannot overlap much and are thus easy to color while avoiding a monochromatic copy. If on the other hand there is a very high number of copies of F on each edge, then these must overlap heavily and coloring without a monochromatic copy is no longer possible.

There are two exceptional cases in Theorem 5.1: stars and paths of length 3. In the case of a star S_ℓ on ℓ edges it is easy to see by the pigeon hole principle that $S_{r(\ell-1)+1} \to (S_\ell)^e_r$ for any $r \geq 2$. In other words, as soon as a star on $r(\ell-1)+1$ edges appears in $G(n,p)$ it is no longer possible to color it with r colors without monochromatic S_ℓ. The threshold for this event is asymptotically lower than the upper bound of $Cn^{-1/m_2(S_\ell)}$. In the case of P_3 a similar phenomenon occurs. Given any cycle C_ℓ of length $\ell \geq 3$ we obtain a "sunshine graph" $S_\ell^{☼}$ by appending one pending edge to each vertex of C_ℓ. For any odd $\ell \geq 5$ it holds that $S_\ell^{☼} \to (P_3)^e_2$. From standard results it follows that whenever $p = cn^{-1}$ there is a small but constant probability that $G(n,p)$ contains such a sunshine graph. Accordingly the 0-statement in (5.2) cannot be of the same type as (5.3).

The generalization of these results to random k-uniform hypergraphs on n vertices $H^k(n,p)$, in which each possible subset of k vertices forms an edge with probability p independently, was studied by Rödl and Ruciński [RR98]. They conjectured that for hypergraphs the same intuition must hold as for the graph case, and that a monochromatic copy of F appears in every coloring whenever the expected number of copies of F per hyperedge exceeds a large constant. They proved this for the 3-uniform clique on 4 vertices and 2 colors. Together with Schacht they later proved their conjecture for k-partite k-uniform hypergraphs [RRS07]. Recently Friedgut, Rödl and Schacht [FRS10] proved the conjecture for arbitrary k-uniform hypergraphs. Similar results were obtained independently by Conlon and Gowers [CG10].

Theorem 5.2 ([FRS10, CG10]). *Let F be a k-uniform hypergraph with maximum degree at least 2, and let $r \geq 2$. There exists a constant $C > 0$ such that for $p = p(n)$ satisfying $p \geq Cn^{-1/m_k(F)}$ we have*

$$\lim_{n\to\infty} \mathbb{P}[H^k(n,p) \to (F)^e_r] = 1.$$

We are not aware of any results which prove a corresponding 0-statement. The above bound is however widely believed to be tight in most cases, with some exceptions similar to those in Theorem 5.1. In this chapter we confirm that this is true for a large class of hypergraphs. However we also show that the list of exceptions is larger than in the graph case, and that in at least one case for which the above is not tight, the bound is given by the so-called *asymmetric Ramsey problem.*

Instead of avoiding a monochromatic copy of the same hypergraph F in all colors as above, in the asymmetric Ramsey case we want to avoid a hypergraph F_1 in red, a hypergraph F_2 in blue, and so on for all $r \geq 2$ colors. Similarly to the classical Ramsey case, we denote the fact that all edge colorings of a graph G contain at least one monochromatic copy of F_i in its respective color as $G \to (F_1, \ldots, F_r)$. Clearly, if all F_i are equal this reduces to the (symmetric) Ramsey case.

This problem was first introduced by Kohayakawa and Kreuter in [KK97], where they determined the threshold for the appearance of some combinations of cycles (and some more general cases). They also conjectured that in the general case the threshold is determined by the function below in the same sense that $m_k(\cdot)$ determines the threshold for the symmetric Ramsey problem. Here we state the extension for k-uniform hypergraphs, the original conjecture concerns only the case $k = 2$.

Definition 5.3. *Let G_1, G_2 be two k-uniform hypergraphs with at least one edge and such that $m_k(G_1) \geq m_k(G_2)$. Then define*

$$m_k(G_1, G_2) = \max\Big\{\frac{e(G_1')}{v(G_1') - k + 1/m_k(G_2)} \mid G_1' \subseteq G_1, e(G_1') \geq 1\Big\}. \tag{5.4}$$

Note that if $G_1 = G_2$, then $m_k(G_1, G_2) = m_k(G_1)$. We say that G_1 is strictly balanced with respect to $m_k(\cdot, G_2)$ if no strict subgraph $G_1' \subsetneq G_1$ with at least one edge maximizes (5.4).

While the above function seems only to apply to the case $r = 2$, it turns out that for larger values of r only the two graphs with maximal k-density fully determine the threshold. Progress towards proving the conjecture in the graph case was made by Marciniszyn et al. in [MSSS09], where they confirmed it for cliques. They also proved a general 1-statement, provided that the so-called KŁR-conjecture holds. Kohayakawa et al. [KSS] then proved the conjectured upper bound for two graphs G and H satisfying some mild conditions. Recently Balogh et al. [BMS12a], and independently Saxton and Thomason [ST12], proved the KŁR conjecture. By the results in [MSSS09] this implies a 1-statement for r many 2-balanced graphs, where the one of maximal 2-density has to be strictly 2-balanced.

Recently Nenadov [Nen13] extended the result by Kohayakawa et al. in

[KSS] to k-uniform hypergraphs, exploiting a containment theorem by Saxton and Thomason [ST12].

Theorem 5.4 ([Nen13])**.** *Let $G_1, \dots, G_r$ be k-uniform hypergraphs such that $m_k(G_1) \geq m_k(G_2) \geq \cdots \geq m_k(G_r)$ and such that G_1 is strictly balanced with respect to $m_k(\cdot, G_2)$. Then there exists a constant $C > 0$ such that for $p = p(n) \geq Cn^{-1/m_k(G_1,G_2)}$*

$$\lim_{n\to\infty} \mathbb{P}[H^k(n,p) \to (G_1, \dots, G_r)] = 1.$$

There are no results proving a matching 0-statement (except the very limited case of Theorem 5.9 below). However due to the similarity of the graph and hypergraph case it would be very surprising if this is not the correct threshold in for most hypergraphs.

5.2 Our results

For any hypergraph F, let $m(F)$ denote the usual density measure

$$m(F) := \max_{\substack{H\subseteq F\\ v(H)\geq 1}} \frac{e(H)}{v(H)}.$$

With standard arguments it follows that Bollobás' small subgraphs theorem [Bol81] extends naturally to hypergraphs as well, i.e. it holds for any hypergraph F that

$$\lim_{n\to\infty} \mathbb{P}[H^k(n,p) \text{ contains a copy of } F] = \begin{cases} 0 & \text{if } p \ll n^{-1/m(F)} \\ 1 & \text{if } p \gg n^{-1/m(F)}. \end{cases} \tag{5.5}$$

Our main result shows that Theorem 5.2 is indeed tight in many cases, but that in contrast to the graph case the list of exceptions is much larger. Because of this we resort to the following definition.

Definition 5.5 (Ramsey-density-obeying)**.** *Let F be a k-uniform hypergraph, and let $r \geq 2$. We say that F is* Ramsey-density-obeying *for r colors if all hypergraphs H satisfy*

$$m(H) \leq m_k(F) \Rightarrow H \nrightarrow (F)^e_r. \tag{5.6}$$

This condition is clearly necessary for a matching 0-statement to hold. If a graph F is not Ramsey-density-obeying, then by (5.6) the threshold for the appearance of a copy of some H with $m(H) \leq m_k(F)$ and $H \to (F)^e_r$ is an upper bound on the 0-statement.

Note that in the case of graphs there are only 2 examples of not Ramsey-density-obeying graphs: forests of stars, and forests of stars and at least one P_3, cf. parts i) and ii) of Theorem 5.1.

Our main result is a 0-statement which matches the 1-statement in Theorem 5.2 for a large class of graphs.

Theorem 5.6. *Let F be a k-uniform hypergraph containing a subgraph $H \subseteq F$ which is strictly k-balanced, satisfies $m_k(H) = m_k(F)$, and which is Ramsey-density-obeying. Then there exist positive constants $c = c(F)$, $C = C(F, r)$ such that*

$$\lim_{n\to\infty} \mathbb{P}[H^k(n,p) \to (F)^e_r] = \begin{cases} 0 & p \leq cn^{-1/m_k(F)} \\ 1 & p \geq Cn^{-1/m_k(F)}. \end{cases}$$

Note that we place the condition of being Ramsey-density-obeying on a strictly k-balanced subgraph of F. This is not an artefact of our proof and is a necessary condition for the theorem to hold. At the end of this section we present an example of a Ramsey-density-obeying hypergraph F which does not have a subgraph H as required in the theorem, and which has a strictly lower threshold.

As a straightforward consequence of our proof we also obtain that we can strengthen the definition of Ramsey-density-obeying as follows.

Definition 5.7 (Ramsey-density-obeying, alternative definition)**.** *Let F be a k-uniform hypergraph and $r \geq 2$. We say that F is* Ramsey-density-obeying *for r colors if all hypergraphs H with $v(H) \leq 4v(F)^3/\alpha + 4v(F)^2$, where*

$$\alpha := \frac{1}{2} \min_{\substack{G \subsetneq F \\ v(G) > k+1}} \frac{e(F) - e(G)}{m_k(F)} - \big(v(F) - v(G)\big),$$

satisfy

$$m(H) \leq m_k(F) \Rightarrow H \nrightarrow (F)^e_r.$$

In other words verifying whether our theorem is applicable or not is in the worst case a finite enumeration problem over all graphs with size at most some constant depending only on F.

We show that all complete hypergraphs are Ramsey-density-obeying, and obtain the following corollary.

Corollary 5.8. *Let K_ℓ^k be the complete k-uniform hypergraph on $\ell > k$ vertices and $r \geq 2$. Then there exist positive constants c, C such that*

$$\lim_{n\to\infty} \mathbb{P}[H^k(n,p) \to (G)_r^e] = \begin{cases} 0 & p \leq cn^{-(\ell-k)/(\binom{\ell}{k}-1)} = cn^{-1/m_k(K_\ell^k)} \\ 1 & p \geq Cn^{-(\ell-k)/(\binom{\ell}{k}-1)} = Cn^{-1/m_k(K_\ell^k)}. \end{cases}$$

It is straightforward to check that both the case of stars and paths of length 3 can be generalized to arbitrary k-uniform hypergraphs, and that both cases remain exceptional. Additionally to this we present a construction which generates for large enough k a *not* Ramsey-density-obeying k-uniform hypergraph out of every (2-uniform) graph.

Let $G = (V, E)$ be a graph and W a set of $k-2$ additional vertices with $V \cap W = \emptyset$. We denote by $G^{+(k-2)} = (V', E')$ the k-uniform hypergraph obtained by adding the vertices of W to each edge of G, i.e. we set $V' = V \cup W$ and $E' = \{e \cup W \mid e \in E\}$. If the dependency on k is clear from the context we write G^+ instead of $G^{+(k-2)}$.

As a first example consider the complete graphs K_3 and K_6. It holds that $K_6 \to (K_3)_2^e$. It is easy to see that the same holds for K_3^+ and K_6^+ for any given uniformity $k \geq 3$. It is also easy to check that $m_k(K_3^+) = 2$ regardless of k, while $m(K_6^+) = 15/(4+k)$ is strictly decreasing in k. For $k = 4$ we obtain $m_k(K_3^+) = 2$ and $m(K_6^+) = 15/(6+k-2) = 15/8 < 2$, i.e. K_3^{+2} is not Ramsey-density-obeying.

This construction can be generalized for any graph G with at least 2 edges. By Ramsey's theorem for any graph G there exists some ℓ such that $K_\ell \to (G)_r^e$. The value of $m_k(G^+)$ remains constant for all $k \geq 2$, while for k large enough it holds that $m_k(G^+) > m(K_\ell^+)$. This however implies that G^+ is not Ramsey-density-obeying for k large enough.

In the graph case there are two possible "types" of thresholds for the Ramsey problem: either they are exceptional, and essentially due to the appearance of a small counter-example (cases i) and ii) in Theorem 5.1), or they behave "nicely". In the case of hypergraphs the picture is more complex, as the following example of a 7-uniform hypergraph demonstrates.

Denote by $C_{6,20}$ the 7-uniform cycle with 20 vertices and 20 edges such that two successive edges overlap in exactly 6 vertices, by C_4 the graph cycle on 4 edges, and by C_4^+ the corresponding 7-uniform hypergraph obtained by the construction above.

Theorem 5.9. *Let $F = C_4^+ \uplus C_{6,20}$ be the disjoint union of a copy of C_4^+ and $C_{6,20}$. Then there exist positive constants c and C such that*

$$\lim_{n\to\infty} \mathbb{P}[H^7(n,p) \to (F)_2^e] = \begin{cases} 0 & p \le cn^{-1/m_7(C_4^+, C_{6,20})} \\ 1 & p \ge Cn^{-1/m_7(C_4^+, C_{6,20})}, \end{cases}$$

where $m_7(C_4^+, C_{6,20})$ is defined in (5.4) *and satisfies*

$$m_7(C_4^+) > m_7(C_4^+, C_{6,20}) > m_7(C_{6,20})$$

The title page of this thesis shows a picture of $C_4^+ \uplus C_{6,20}$.

The hypergraph $C_4^+ \uplus C_{6,20}$ is Ramsey-density-obeying, but the only maximal strictly k-balanced subgraph is C_4^+, which is not Ramsey-density-obeying. We fully explore only this example, however we believe that this type of construction would turn up many more similar examples.

Note that in the graph case (i.e. $k = 2$) the only strictly 2-balanced subgraph of any tree is the star with 2 edges. This subgraph is however not Ramsey-density-obeying (it is a star) and thus no trees are covered by our main theorem. By Theorem 5.1 trees T which are not stars or paths on 3 edges are however not exceptional and do have a threshold at $p = \Theta(n^{-1/m_2(T)})$. Our theorem is therefore not complete. At present we have no characterization which distinguishes hypergraphs which are not covered by our theorem but which do have a 0-statement matching the 1-statement in Theorem 5.2, and those for which the 1-statement is not tight, such as the above example of $C_4^+ \uplus C_{6,20}$.

5.3 Proof of the main result

In this section we prove Theorem 5.6. Note that the 1-statement is given by Theorem 5.2, so we only need to prove the 0-statement. We split the proof of the 0-statement in two parts. In this section we present a probabilistic proof of the 0-statement which makes use of a key deterministic lemma. We give the proof of this lemma in the next section.

The ideas used here are an evolution of some ideas already present in [RR93] and [MSSS09]. We applied a similar technique in [GPST12] as well.

For this section we assume that the number $r \ge 2$ of colors and the forbidden hypergraph F are fixed. We prove that we can 2-color $H^k(n,p)$ for

$p \leq cn^{-1/m_k(F)}$ without a monochromatic copy of F. We can assume without loss of generality that F is strictly k-balanced and Ramsey-density-obeying. If this is not the case, then we replace F by a hypergraph H as in the claim. Proving that we can 2-color $H^k(n,p)$ without a copy of H implies the claim for F.

We may also assume that F contains at least 3 edges. The case of 1 edge is trivial. All strictly k-balanced hypergraphs F' on 2 edges are such that the two edges intersect in ℓ common vertices, $1 \leq \ell \leq k-1$, and it holds $m_k(F') = 1/(k-\ell)$. The hypergraph F^* formed by $r+1$ edges all intersecting in ℓ common vertices cannot be colored without a monochromatic copy of F', but it holds that $m(F^*) < m_k(F')$. It follows that no strictly k-balanced hypergraph with 2 edges is Ramsey-density-obeying.

We need to show that for $p \leq cn^{-1/m_k(F)}$ it holds that

$$\lim_{n\to\infty} \mathbb{P}[H^k(n,p) \to (F)^e_r] = 0.$$

For any hypergraph H we call edges of H which are not contained in two otherwise edge-disjoint copies of F *open*. We call all other edges *closed*.

The reason for this distinction is that open edges are easy to color. Assume there exists a valid r-coloring (i.e. one without monochromatic copy of F) for $H-e$, the hypergraph obtained from H by removing e from the edge set. Then we can trivially extend this coloring to H itself by using the fact that e is open. We can always assign either the color red or blue to e unless e is contained in two copies of F which are (up to e) already monochromatic red and blue. This implies that these two copies are edge-disjoint and contradicts the fact that e is open. We can therefore find a valid r-coloring of $H^k(n,p)$ by running Algorithm 1.

$\hat{H} := H^k(n,p)$
while $\hat{H}$ *contains an open edge* e **do**
  $\hat{H} \leftarrow \hat{H} - e$
end
Color $\hat{H}$
Add the removed edges in reverse order and color them appropriately.

Algorithm 1: Coloring $H^k(n,p)$.

Of course, the step "Color $\hat{H}$" is the difficult one. Next we show that this is always possible.

Note that $\hat{H}$ has the property that every edge is closed, i.e. contained in two otherwise edge disjoint copies of F. We call it the *F-core* of $H^k(n,p)$. Also note that the order in which the open edges are removed is irrelevant and that the F-core of $H^k(n,p)$ is uniquely defined. We call a subgraph H of $\hat{H}$ *F-closed* if every copy of F is either fully contained in H or edge disjoint with it. It is easy to see that we can partition the F-core into minimal (by subgraph inclusion) F-closed subgraphs which can all be colored independently. Assume that we find a valid coloring for all minimal F-closed subgraphs, then their union yields a valid coloring for $\hat{H}$ and therefore with the above algorithm for all of $H^k(n,p)$. The proof of Theorem 5.6 now follows easily from the next lemma.

Lemma 5.10. *Let F be a strictly k-balanced hypergraph with at least 3 edges. There exist constants $c = c(F) > 0$ and $L = L(F) > 0$ such that if $p \leq cn^{-1/m_k(F)}$ then a.a.s. every minimal F-closed subgraph of $H^k(n,p)$ has size at most L.*

Proof of Theorem 5.6. Let $c = c(F)$ and $L = L(F)$ be as in Lemma 5.10. Let H be any hypergraph with $v(H) \leq L$ and $e(H) > v(H) \cdot m_k(F)$. For $p \leq cn^{-1/m_k(F)}$ the expected number of copies of H in $H^k(n,p)$ is $\Theta(n^{v(H)}p^{e(H)}) = o(1)$. As there can be at most a constant number (depending only on L) of such hypergraphs H, it follows by Markov's inequality that $H^k(n,p)$ a.a.s. contains *no* subgraph H of at most L vertices and $m(H) > m_k(F)$. This implies in particular that for any minimal F-closed subgraph H' it holds that $m(H') \leq m_k(F)$. As F is Ramsey-density-obeying, by definition there must exist a valid coloring for H'. By the arguments outlined above this implies the same for the F-core and thus for all of $H^k(n,p)$. □

It remains to show Lemma 5.10. The proof is essentially a first moment argument. We enumerate all possible minimal F-closed hypergraphs of length more than L and show that the probability that one or more of them appears in $H^k(n,p)$ is $o(1)$.

Let G be some minimal F-closed hypergraph. We assume that some arbitrary total ordering on the vertices of G is given. By lexicographic ordering this induces a total ordering on the edges of G as well, i.e. we can always choose a well-defined minimal edge out of any edge set. For the enumeration aspect of the problem we map any minimal F-closed hypergraph G to a sequence of copies of F via Algorithm 2, shown on the next page.

$F_1 \leftarrow$ any copy of F in G
$G_1 \leftarrow F_1$
$i \leftarrow 1$
while $G_i \neq G$ **do**
    $i \leftarrow i + 1$
    if G_{i-1} *contains an open edge* **then**
        $j \leftarrow$ smallest index $j < i$ such that F_j contains open edges
        $e \leftarrow$ the minimal open edge in F_j
        $F_i \leftarrow$ a copy of F in G but not G_{i-1} which contains e
    else
        $F_i \leftarrow$ an arbitrary copy of F in G but not G_{i-1} which intersects G_{i-1} in at least one edge
    end
    $G_i \leftarrow G_{i-1} \cup F_i$
end

Algorithm 2: Decomposing minimal F-closed hypergraphs

It is easy to see that this algorithm is correct and terminates, as every edge of G is contained in a copy of F and G is connected.

Note that the sequence $S := (F_1, \dots, F_\ell)$ fully describes a run of the algorithm. We call it a *grow sequence* for G and each F_i in it a *step* of the sequence, $1 \leq i \leq \ell$. With this we can turn the problem of enumerating all minimal F-closed hypergraphs into the simpler one of enumerating all grow sequences which may appear as output of Algorithm 2.

We fix some arbitrary labeling on the vertices of F. With this we can represent every copy of F in G as an ordered tuple of $v(F)$ vertices and accordingly every grow sequence as an ordered tuple of tuples of $v(F)$ vertices each.

Given some grow sequence $S = (F_1, \dots, F_\ell)$ for G we can easily reconstruct G as the union of all F_i, $1 \leq i \leq \ell$. In fact, we use S and the hypergraph generated from it (which we denote $H(S)$) interchangeably. In particular if we write that S is contained in $H^k(n,p)$, then we mean that all edges of $H(S)$ appear in $H^k(n,p)$. In this section in the context of some step F_i we use the notation G_{i-1} and G_i equivalently to their use in Algorithm 2, i.e. $G_i = H((F_1, \dots, F_i))$.

We show that the expected number of sequences with length ℓ between L (a constant fixed later) and $\ell_{\max} = \Theta(\log n)$ which are contained in $H^k(n,p)$

for $p \leq cn^{-1/m_k(F)}$ is $o(1)$. As all sequences longer than $\ell_{\max}$ must have a prefix of length $\ell_{\max}$, the claim follows by Markov's inequality.

The key idea is to bound the number of possible sequences by multiplying for each step F_i the number of choices available for the tuple representing F_i. Each step introduces $e_{\text{new}} \geq 1$ many new edges which are not present in G_{i-1}, and therefore adds a factor of $p^{e_{\text{new}}}$ to the probability that a sequence containing F_i appears in $H^k(n,p)$. As we see later for most steps the result of multiplying the number of choices for F_i times $p^{e_{\text{new}}}$ is bounded by the constant $c < 1$ in Lemma 5.10, and in some cases even by some function in $o(1)$. In other words, the longer the sequence, the less likely it is to appear in $H^k(n,p)$.

We distinguish various different step types. We call F_1 the *first* step. For $i \geq 2$ we call steps in which $H = G_{i-1} \cap F_i$ corresponds to exactly one edge *regular*, and call all other steps *degenerate*. Further we say that a step is *open* or *closed* depending on whether F_i was chosen in Line 9 or Line 11 of Algorithm 2.

In the following we discuss the number of choices available for each step type. For this we assume that a sequence S of length ℓ is encoded as ℓ ordered tuples of $v(F)$ many vertices of $H^k(n,p)$.

For an open regular step F_i the intersection with G_{i-1} corresponds exactly to the minimal open edge in the lowest-index F_j with open edges in G_{i-1}. This edge and its vertices are therefore fixed by all preceding steps, up to its exact embedding into F_i. It remains to choose the other $v(F) - k$ new vertices and the $e(F) - 1$ new edges. The total contribution of such a step to our bound is

$$v(F)^k n^{v(F)-k} p^{e(F)-1} \leq v(F)^k c^{e(F)-1} \leq c, \tag{5.7}$$

where c is the constant in Lemma 5.10 which we choose small enough for the above to hold (and such that $\log(c) < -1$, which we need later).

In contrast to open regular steps, if a step F_i is closed regular the intersection of F_i and G_{i-1} is not fully determined by G_{i-1} and we need to choose it. At step i the hypergraph G_{i-1} contains at most $v(F) \cdot i$ many vertices, so we obtain

$$v(F)^k (v(F) \cdot i)^k n^{v(F)-k} p^{e(F)-1} \leq c \cdot (v(F) \cdot i)^k.$$

Now consider the case of degenerate steps, i.e. those for which $H := F_i \cap G_{i-1}$ satisfies $v(H) > k$. Recall that F is strictly k-balanced, so for any subgraph $H \subsetneq F$ with $v(H) > k$ we have

$$\frac{e(H)-1}{v(H)-k} < \frac{e(F)-1}{v(F)-k} = m_k(F),$$

and thus

$$\frac{e(F)-e(H)}{v(F)-v(H)} = \frac{(e(F)-1)-(e(H)-1)}{(v(F)-k)-(v(H)-k)} > m_k(F). \tag{5.8}$$

This implies that we can choose some constant $\alpha > 0$ such that regardless of the choice of H with $v(H) > k$ it holds that

$$v(F) - v(H) - \frac{e(F)-e(H)}{m_k(F)} < -\alpha.$$

Applying this to a (open or closed) degenerate step F_i we obtain

$$\sum_{\substack{H \subseteq F \\ v(H)>k}} (v(F)\cdot i)^{v(H)} n^{v(F)-v(H)} p^{e(F)-e(H)} \leq (v(F)\cdot i)^{v(F)} n^{-\alpha}. \tag{5.9}$$

Equations (5.7)-(5.9) show that closed regular steps make sequences containing them more likely to appear, especially if they appear late in the sequence, while open regular steps make them slightly less likely. Degenerate steps on the other hand introduce a factor $n^{-\alpha}$ which suggests that sequences containing many of them are very unlikely to appear in $H^k(n,p)$. The next lemma gives some crucial bounds on the number of degenerate and regular closed steps.

Lemma 5.11. *Let S be a grow sequence of length ℓ for some minimal F-closed hypergraph.*

i) If S contains at most d degenerate steps, then $\ell < 4d \cdot v(F) + 1$.

ii) If a prefix S_i of S, $1 \leq i \leq \ell$, contains at most d degenerate steps, then S_i contains no closed regular steps F_j with $j > 4d \cdot v(F) + 1$.

The proof of this lemma is deferred to the next section.

We can now finish our first moment argument. Set $d_{\max} := v(F)/\alpha + 1$ and $L' = 4d_{\max}v(F) = 4v(F)^2/\alpha + 4v(F)$. By Lemma 5.11 all sequences longer than L' must contain at least $d_{\max}$ degenerate steps. Set $\ell_{\max} :=$

$v(F)\log(n) + d_{\max} + 1$. We consider two cases: the first are those sequences which have their $d_{\max}$th degenerate step before $\ell_{\max}$. We truncate them after the $d_{\max}$th step. The second are those sequences whose $d_{\max}$th step appears after $\ell_{\max}$. We truncate these at length $\ell_{\max}$. Note that by Lemma 5.11 in both these cases closed regular steps can only happen in the first L' steps of the sequence.

In the first case we obtain

$$\sum_{\ell=L'+1}^{\ell_{\max}-1} n^{v(F)} \tbinom{\ell}{d_{\max}} \big((v(F)\cdot \ell)^{v(F)} n^{-\alpha}\big)^{d_{\max}} \big((v(F)\cdot L')^k\big)^{L'} =$$
$$= \mathcal{O}(\operatorname{polylog}(n)\cdot n^{v(F)} n^{-\alpha\cdot d_{\max}}) = o(1).$$

Here we bound the contribution of the first step by $n^{v(F)}$, drop the contribution of $c < 1$ for all regular steps, and use the fact that only the first L' steps may be closed regular.

Recall that $\log(c) < -1$. In the second case we thus obtain

$$\sum_{d=0}^{d_{\max}} n^{v(F)} \tbinom{\ell_{\max}}{d} \big((v(F)\cdot \ell_{\max})^{v(F)} n^{-\alpha}\big)^{d} \big((v(F)\cdot L')^k\big)^{L'} c^{\ell_{\max}-d-1}$$
$$= \mathcal{O}(\operatorname{polylog}(n)\cdot n^{v(F)} \cdot c^{\ell_{\max}-d-1})$$
$$= \mathcal{O}(\operatorname{polylog}(n)\cdot n^{v(F)(1+\log c)}) = o(1).$$

As a sequence of length L' can have at most $L := v(F)\cdot L'$ many vertices this finishes the proof of Lemma 5.10 and thus of Theorem 5.6.

Note that while in this section we take a decidedly probabilistic view, the above proof can also be used to yield a deterministic statement concerning the $m(\cdot)$-density of not Ramsey-density-obeying hypergraphs. This is also the reason for our alternative definition of Ramsey-density-obeying in Definition 5.7.

5.4 Proof of the bound on sequence lengths

Let $S = (F_1, \ldots, F_\ell)$ be some grow sequence. For all $1 \leq i \leq \ell$, let $S_i := (F_1, \ldots, F_i)$. For any S_i and any regular step F_j, $j \leq i$ (considered as a copy of F) we call the edge $e := E(G_{i-1}) \cap E(F_j)$ the *attachment edge* of F_j and the vertices in $V(F_j) \setminus V(G_{j-1})$ the *inner vertices* of F_j. We

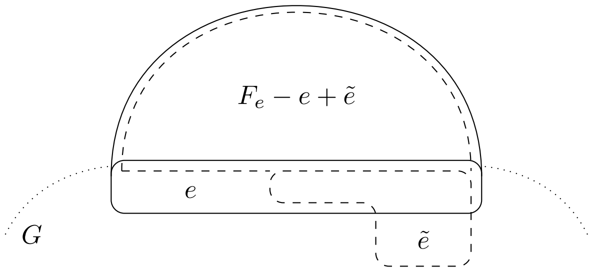


Figure 5.1: The possible copies of F created in a regular step. The solid lines represent F_e, the dashed ones $\tilde{F}$.

say that F_j is *fully open* in S_i if no step $F_{j'}$, $j < j' \leq i$ contains an inner vertex of F_j and every edge in F_j except e is open in S_i. Lastly we denote by $\mathrm{fo}(S_i)$, $\mathrm{reg}(S_i)$ and $\deg(S_i)$ the number of regular, degenerate and fully open steps in S_i. We also consider the first step to be fully open in S_1.

This first lemma implies that newly added regular steps F_i are fully open in S_i (and that in the definition above the requirement that all edges be open is redundant).

Lemma 5.12. *Let G be an arbitrary hypergraph and let F_e be a copy of a strictly k-balanced hypergraph F with at least 3 edges which intersects G in exactly one edge $e \in E(G)$. Further let*

$$\gamma := \max\{|e_1 \cap e_2| \mid e_1, e_2 \in E(F) \wedge e_1 \neq e_2\}$$

be the maximum number of vertices in the intersection of two edges of F. Then all copies $\tilde{F}$ of F in $G^+ = G \cup F_e$ which are not contained in G have the form

$$\tilde{F} = F_e - e + \tilde{e} := \big((V(F_e) \setminus e) \cup \tilde{e}, (E(F_e) \setminus \{e\}) \cup \{\tilde{e}\}\big),$$

where $\tilde{e} \in E(G)$ and $|\tilde{e} \cap e| > \gamma$, cf. Figure 5.1

Proof. Let $\tilde{F}$ be some copy of F in G^+ which is not fully contained in G. Note that if $\tilde{F} = F_e$ then the lemma is true for $\tilde{e} = e$, so we assume $\tilde{F} \neq F_e$.

Let $\tilde{e}$ be an arbitrary edge of $\tilde{F}$ which is not contained in $E(F_e)$. Note that this implies $\tilde{e} \in E(G)$.

First we show that $E(\tilde{F}) \setminus \{\tilde{e}\}$ must be contained in $E(F_e) \setminus \{e\}$, which implies that the two sets are equal. Assume this is not true. Set $\tilde{F}_{\text{new}} := \tilde{F}[V(F_e)]$, $\tilde{F}_{\text{old}} := \tilde{F}[V(G)]$ and $\tilde{F}^{+e}_{\text{new}} = \tilde{F}_{\text{new}} \cup e$. We assumed that $E(\tilde{F}) \setminus \{\tilde{e}\} \not\subseteq E(F_e) \setminus \{e\}$, and therefore $\tilde{F}$ must contain some edge different from $\tilde{e}$ which is not contained in $E(F)$, and is thus contained in $E(G)$. This implies that $e(\tilde{F}_{\text{old}}) \geq 2$. As $\tilde{F}$ is not fully contained in G it must contain at least one edge of $E(F_e) \setminus E(G)$. As $e \in E(G)$ and $e \in E(\tilde{F}^{+e}_{\text{new}})$ it also holds that $e(\tilde{F}^{+e}_{\text{new}}) \geq 2$. The intersection of $\tilde{F}_{\text{old}}$ and $\tilde{F}_{\text{new}}$ can contain at most one edge (i.e. e) so both $\tilde{F}_{\text{old}}$ and $\tilde{F}^{+e}_{\text{new}}$ must be strict subgraphs of F.

Regardless of whether e was already an edge of $\tilde{F}_{\text{new}}$ or not it holds that

$$e(\tilde{F}) = e(\tilde{F}_{\text{old}}) + e(\tilde{F}^{+e}_{\text{new}}) - 1 \quad \text{and} \quad v(\tilde{F}) \geq v(\tilde{F}_{\text{old}}) + v(\tilde{F}^{+e}_{\text{new}}) - k.$$

We thus have

$$m_k(F) = m_k(\tilde{F}) = \frac{e(\tilde{F}) - 1}{v(\tilde{F}) - k} \leq \frac{e(\tilde{F}_{\text{old}}) - 1 + e(\tilde{F}^{+e}_{\text{new}}) - 1}{v(\tilde{F}_{\text{old}}) - k + v(\tilde{F}^{+e}_{\text{new}}) - k} < m_k(F),$$

which is a contradiction. Here the last inequality follows from the fact that F is strictly k-balanced, that $\tilde{F}^{+e}_{\text{new}} \neq F$ and $\tilde{F}_{\text{old}} \neq F$, and Proposition 2.1.

It remains to establish that $|\tilde{e} \cap e| > \gamma$. Assume this is not the case, then the graph $F_{e,\tilde{e}} := \big(e \cup \tilde{e}, \{e, \tilde{e}\}\big)$ obtained by the union of e and $\tilde{e}$ is a strict subgraph of F. The same holds for $\tilde{F}_{\text{new}} = \tilde{F}[V(F_e)]$. It is easy to check that $e(\tilde{F}) = e(\tilde{F}_{\text{new}}) + e(F_{e,\tilde{e}}) - 1$ and $v(\tilde{F}) = v(\tilde{F}_{\text{new}}) + v(F_{e,\tilde{e}}) - k$. Therefore we obtain

$$m_k(F) = m_k(\tilde{F}) = \frac{e(\tilde{F}) - 1}{v(\tilde{F}) - k} = \frac{e(\tilde{F}_{e,\tilde{e}}) - 1 + e(\tilde{F}_{\text{new}}) - 1}{v(\tilde{F}_{e,\tilde{e}}) - k + v(\tilde{F}_{\text{new}}) - k} < m_k(F).$$

The last inequality holds by Proposition 2.1 and the fact that both $F_{e,\tilde{e}}$ and $\tilde{F}_{\text{new}}$ are strict subgraphs of the strictly k-balanced hypergraph F. This is again a contradiction. □

Corollary 5.13. *Let $S = (F_1, \ldots, F_\ell)$ be some grow sequence. For all $i \leq \ell$ and all regular steps F_j, $j \leq i$ it holds that if no step $F_{j'}$ with $j < j' \leq i$ contains an inner vertex of F_j, then F_j is fully open in S_i.*

Proof. Let F_j be a regular step as in the claim and let f_j be its attachment edge. As F contains at least 3 edges, for every edge e in $E(F_j) \setminus \{f_j\}$ the same set contains a second edge e' different from both e and f_i. By Lemma 5.12 any two copies of F in $H(S_i)$ which overlap in e must also overlap in e', and thus e cannot be closed. □

For $i \geq 1$ let $\kappa(i)$ denote the number of fully open copies "destroyed" by step i, i.e. let

$$\kappa(i) = |\{j < i \mid F_j \text{ fully open in } S_{i-1} \text{ but not } S_i\}|.$$

If F_i is a regular step, then clearly $\kappa(i) \leq 1$, while if it is a degenerate step then $\kappa(i) \leq v(F) - k + 1$. The reason for this is that a degenerate step must intersect G_{i-1} in at least one pre-existing edge. This edge intersects the inner vertices of at most one fully open step.

Lemma 5.14. *Let $F_i, \ldots, F_{i+e(F)-2}$ be a sequence of consecutive regular steps such that $\kappa(i) = 1$. Then $\kappa(i+1) = \cdots = \kappa(i + e(F) - 2) = 0$.*

Proof. As $\kappa(i) = 1$ it holds that F_i is the first step which intersects the inner vertices of some fully open step F_j, $j < i$. Before step F_i by definition F_j has $e(F) - 1$ open edges. We show that F_i and the $e(F) - 2$ following regular steps each only close exactly one open edge of F_j. This implies that before each step $F_{i+\ell}$, $1 \leq \ell \leq e(F) - 2$, the step F_j still has open edges, and thus by our grow sequence algorithm $F_{i+\ell}$ intersects one of these open edges. As $F_{i+\ell}$ is regular it does not intersect the inner vertices of any other fully open step and thus $\kappa(i + \ell) = 0$.

Assume some $F_{i+\ell}$, $0 \leq \ell \leq e(F) - 2$, closes more than just its attachment edge $f_{i+\ell}$. Then there is some other edge $\tilde{e} \in E(F_j)$ which after step $F_{i+\ell}$ becomes part of a copy of F otherwise edge-disjoint with F_j. However by Lemma 5.12 any such copy is of the form $F_{i+\ell} - f_{i+\ell} + \tilde{e}$ and $|f_{i+\ell} \cap \tilde{e}| > \gamma$. As both $f_{i+\ell}$ and $\tilde{e}$ were added together in step F_j it must hold that $|f_{i+\ell} \cap \tilde{e}| \leq \gamma$ which is a contradiction. □

Now we prove a lower bound on the number of fully open copies of F that can be contained in any grow sequence of length ℓ and with at most d degenerate steps.

Lemma 5.15. *Let S be a grow sequence of length ℓ. For all $1 \leq i \leq \ell$ it holds that*

$$\mathrm{fo}(S_i) \geq \mathrm{reg}(S_i) \cdot \left(1 - \frac{1}{e(F) - 1}\right) - \deg(S_i) \cdot v(F). \tag{5.10}$$

Proof. Let $\varphi(i) := \mathrm{fo}(S_i) - \mathrm{reg}(S_i) \cdot \left(1 - \frac{1}{e(F)-1}\right) + \deg(S_i) \cdot v(F)$. To prove the lemma we need to show that $\varphi(i) \geq 0$ for all $i \leq \ell$. In fact we prove the following stronger statement

$$\varphi(\ell) \geq \begin{cases} 1 & \text{if } F_i \text{ is degenerate} \\ 0 & \text{otherwise.} \end{cases}$$

We prove this by induction. For $i = 1$ we have $\mathrm{reg}(S_1) = \deg(S_1) = 0$ but $\mathrm{fo}(S_1) = 1$. Assume now that step F_i, $i \leq \ell$, is degenerate. Then $\varphi(i) - \varphi(i-1) = v(F) - \kappa(i) > 1$ and the claim follows. If F_i is regular, let $j < i$ be the largest index of all previous steps with $\kappa(j) > 0$ or for degenerate steps even $\kappa(j) = 0$. (If both do not exist set $j = 0$.) It holds that $\varphi(i) - \varphi(j) = (i-j)/(e(F)-1) - \kappa(i)$. If $\kappa(i) = 0$ or if $\varphi(j) \geq 1$ (F_j is degenerate) we are done. If that is not the case then by Lemma 5.14 $i - j \geq e(F) - 1$, and thus $\varphi(i) \geq 0$. □

Now we are ready to prove Lemma 5.11.

Proof of Lemma 5.11. We prove part i) first. Let S be a grow sequence of length ℓ with at most d degenerate steps. The hypergraph $H(S)$ contains no open edges by definition, and thus also no fully open steps. By Lemma 5.15 this implies that

$$\deg(S) \cdot v(F) \geq \mathrm{reg}(S) \cdot \left(1 - \frac{1}{e(F)-1}\right)$$

must hold. We have $\deg(S) \geq d$ and $\mathrm{reg}(S) \leq \ell - d - 1$ (the first step is neither degenerate nor regular). We obtain

$$d \cdot v(F) \geq (\ell - d - 1)\left(1 - \frac{1}{e(F)-1}\right).$$

Solving for ℓ and using $1 - 1/(e(F)-1) \geq 1/2$ we obtain

$$\ell \leq 2d\big(v(F) + 1\big) + 1 < 4d \cdot v(F) + 1,$$

which proves the first part.

For the second part of Lemma 5.11 let S_i be some prefix of S, $1 \leq i \leq \ell$ with at most d degenerate steps. Note that before any closed regular step $j \leq i$, $H(S_{j-1})$ contains no open edges. By the above inequality it must therefore hold that $j - 1 < 4d \cdot v(F) + 1$ and the claim follows. □

5.5 Complete hypergraphs

In this section we prove that all complete k-uniform hypergraphs K_ℓ^k on $\ell > k$ vertices satisfy the assumptions of Theorem 5.6 (and Corollary 5.8 thus holds). The fact that K_ℓ^k is strictly k-balanced is easy to check, so we only have prove that K_ℓ^k is Ramsey-density-obeying.

A hypergraph G is called *ℓ-degenerate* if $\ell \geq \delta(H)$ for all $H \subseteq G$.

Lemma 5.16. *Every k-uniform hypergraph G is $\lfloor k \cdot m(G) \rfloor$-degenerate.*

Proof. Let $H \subseteq G$ be any subgraph of G. Then

$$\frac{\sum_{v \in V(H)} \deg_H(v)}{v(H)} = \frac{ke(H)}{v(H)} \leq k \cdot m(G).$$

The left hand side of the equation corresponds to the average degree of H. As there must be at least one vertex with at most average degree it easily follows that every subgraph of G must contain at least one vertex with degree at most $k \cdot m(G)$. As degrees are integers this can be improved to $\lfloor k \cdot m(G) \rfloor$. □

Let $r_e(F)$ denote the *size-Ramsey number* of F, that is

$$r_e(F) = \min\{e(G) : G \to (F)_2^e\}$$

and denote by $R(\ell)$ the classical Ramsey number, i.e. the minimum number m such that $K_m \to (K_\ell)_2^e$. We make use of the following theorem by Erdős, Faudree, Rousseau and Schelp [EFRS78].

Theorem 5.17 ([EFRS78])**.** *For all integers $\ell \geq 2$ we have $r_e(K_\ell) = \binom{R(\ell)}{2}$.*

This bound easily generalizes to hypergraphs by using the notion of a link. Let H be a k-uniform hypergraph, v a vertex in H and H_v the subgraph of H induced by the edges incident to v. The *link* of v is the $(k-1)$-uniform hypergraph obtained by removing v from every edge of H_v.

Corollary 5.18. *Let all integers $\ell > k \geq 2$ we have*

$$r_e(K_\ell^k) \geq \binom{R(\ell - k + 2) + k - 2}{k}.$$

Proof. We proceed by induction on $k \geq 2$. The base case $k = 2$ immediately follows from Theorem 5.17. So assume $k \geq 3$ and let H be a k-uniform hypergraph with $H \to (K_\ell^k)_2^e$ and $e(H) = r_e(K_\ell^k)$. We claim that

$$\delta(H) \geq r_e(K_{\ell-1}^{k-1}). \tag{5.11}$$

Note that otherwise there exists $v \in V(H)$ such that the link of v is a $(k-1)$-uniform hypergraph with less then $r_e(K_{\ell-1}^{k-1})$ edges, that is we can color the edges of the link of v without a monochromatic copy of $K_{\ell-1}^{k-1}$. Moreover since $e(H) = r_e(K_\ell^k)$ we know that there exists a valid edge-coloring of $H - v$ without a monochromatic copy of K_ℓ^k. The edge-coloring of the link of v then extends this coloring to one of H, which is a contradiction.

Hence, (5.11) holds and with the induction hypothesis we obtain

$$\delta(H) \geq r_e(K_{\ell-1}^{k-1}) \geq \binom{R(\ell-k+2)+k-3}{k-1}.$$

Clearly, the minimum degree can only be that large if $H - v$ contains at least $R(\ell - k + 2) + k - 3$ vertices. As $e(H) \geq v(H)\delta(H)/k$ this implies

$$e(H) \geq \frac{R(\ell-k+2)+k-2}{k}\binom{R(\ell-k+2)+k-3}{k-1} = \binom{R(\ell-k+2)+k-2}{k}.$$

□

Lemma 5.19. *Let $\ell > k \geq 3$. Then*

$$r_e(K_{\ell-1}^{k-1}) > k \cdot m_k(K_\ell^k).$$

Proof. In the light of Corollary 5.18 we only need to show that

$$\binom{R(\ell-k+2)+k-3}{k-1} > k \cdot m_k(K_\ell^k).$$

Observe that, trivially, $R(n) > (n-1)^2$ (consider, for example, $n-1$ sets of $n-1$ vertices each and color all edges within a set red and edges between sets blue). Using $R(3) = 6$ we deduce $R(n) > n + 2$ for all $n \geq 3$. Thus

$$\binom{R(\ell-k+2)+k-3}{k-1} > \binom{\ell+1}{k-1} = \frac{(\ell+1)\cdot k}{(\ell-k+1)(\ell-k+2)}\binom{\ell}{k} \geq \frac{k\cdot(\ell+1)}{(\ell-k)(\ell-k+5)}\binom{\ell}{k}.$$

As $k \cdot m_k(K_\ell^k) = \frac{k(\binom{\ell}{k}-1)}{\ell-k}$ this completes the proof for $k \geq 4$.

For the case $k = 3$, $\ell \geq 5$ we apply Theorem 5.17 to obtain

$$r_e(K_{\ell-1}^2) = \binom{R(\ell-1)}{2} > \binom{(\ell-2)^2}{2}.$$

It holds that $3 \cdot m_3(K_\ell^3) < \ell^3/(2(\ell-3))$, which is easily verified to be less than the above for all $\ell \geq 5$. The last remaining case, i.e. $k = 3$, $\ell = 4$, is also easily verified by hand using $R(3) = 6$ and Theorem 5.17. $\square$

With these preparations at hand the proof that complete hypergraphs are Ramsey-density-obeying is now easy.

Lemma 5.20. *For all integers $\ell > k \geq 3$ the complete k-uniform hypergraph K_ℓ^k is Ramsey-density-obeying.*

Proof. Let G be a k-uniform hypergraph with density $m(G) \leq m_k(K_\ell^k)$. We need to show that where exists an edge coloring of G with two colors that does not contain a monochromatic copy of K_ℓ^k. By Lemma 5.16 we know that G is $\lfloor k \cdot m_k(K_\ell^k) \rfloor$-degenerate. We can therefore order the vertices of G as $(v_1, \ldots, v_{v(G)})$ such that for each v_i, $1 \leq i \leq v(G)$, we have $\deg_{G_i}(v_i) \leq k \cdot m_k(K_\ell^k)$, where $G_i := G[v_1, \ldots, v_i]$.

We color G by iteratively. Clearly we can 2-color G_1 without a forbidden monochromatic clique. Assume there is a valid 2-coloring for G_{i-1} and consider the vertex v_i. It has degree at most $km_k(K_\ell^k)$ and its link therefore contains at most $km_k(K_\ell^k)$ many edges. By Lemma 5.19 we can find a 2-coloring of it without a monochromatic $K_{\ell-1}^{k-1}$. By coloring each edge incident to v_i in G_i according to the corresponding color in a $K_{\ell-1}^{k-1}$-free 2-coloring of the link we can extend the coloring of G_{i-1} to one of G_i without introducing a forbidden monochromatic K_ℓ^k. $\square$

5.6 The asymmetric example

In this section we prove Theorem 5.9.

Let $F = C_{6,20} \uplus C_4^+$. Recall that

$$m_7(F) = m_7(C_4^+) = 3/2 > m_7(C_4^+, C_{6,20}) = 76/51 > m_7(C_{6,20}) = 19/13.$$

We prove the 1-statement first. Let $p \geq Cn^{-1/m_7(C_4^+, C_{6,20})}$, where C is the constant in Theorem 5.4. Note that C_4^+ and $C_{6,20}$ satisfy the strict balancedness condition in that theorem. It holds that $K_6 \to (C_4)_2^e$ and thus $K_6^+ \to (C_4^+)_2^e$ as well. Further it holds that $m_7(C_4^+, C_{6,20}) > m(K_6^+)$, and thus by (5.5) that $H^7(n,p)$ a.a.s. contains a copy of K_6^+ and therefore a monochromatic copy of C_4^+. As $m_7(C_4^+, C_{6,20}) > m_7(C_{6,20})$ it holds by

Theorem 5.2 that a.a.s. $H^7(n,p)$ also contains a monochromatic copy of $C_{6,20}$. Assuming $H^7(n,p)$ contains such monochromatic copies of C_4^+ and $C_{6,20}$ we pick any two and compare their color. If both have the same color, then we have found a copy of F and are done. If they do not have the same color, assume without loss of generality that the copy of C_4^+ is red and the one of $C_{6,20}$ blue. Then the statement can only be false if there exists an edge coloring of $H^7(n,p)$ which contains no blue copy of C_4^+ and no red copy of $C_{6,20}$. By Theorem 5.4 and our choice of p this is a.a.s. not possible.

It remains to prove the 0-statement. We use a grow-sequences approach very similar to the one in Section 5.3. In this case we say an edge is *closed* if it is contained in both a copy of C_4^+ and a copy of $C_{6,20}$ which are otherwise edge-disjoint, and *open* otherwise. We can again reduce the problem to cores by successively removing open edges from $H^7(n,p)$ until only closed edges remain. We then partition this subgraph further into inclusion-minimal edge-disjoint closed subgraphs. As in the symmetric case we prove that these minimal closed subgraphs are either of a small fixed size depending only on F, or that with high probability they do not appear in $H^7(n,p)$ for p as in the 0-statement.

The definition of a grow sequence cannot be taken over unchanged from Section 5.3, as we need to consider two different graphs simultaneously. This motivates the following definition. Let $\mathcal{C}^*$ be the set of hypergraphs consisting of 1 copy of $C_{6,20}$ with edges $e_0,\dots,e_{19}$ and 19 copies $C_{4,i}^+$ of C_4^+ such that for $1 \le i \le 19$ we have $e_i \in E(C_{4,i}^+)$. The edge e_0 is called the *attachment edge*. Note that these 19 copies of C_4^+ do not need to be distinct as graphs, however the above construction implies a natural matching between edges of $C_{6,20}$ and copies of C_4^+. Pairs consisting of a copy of C_4^+ and the edge of $C_{6,20}$ it is matched to are therefore unique (within any given matching). In the remainder we always assume without loss of generality that for any given $C^* \in \mathcal{C}^*$ the set of 19 copies, the associated matching, and the attachment edge are known; as we can always choose them arbitrarily among all possibilities meeting the definition above.

Note that for every $C^* \in \mathcal{C}^*$ it holds that $v(C^*) \le 58$ and $e(C^*) \le 77$, with equality if and only if all 19 copies of C_4^+ are distinct, intersect the copy of $C_{6,20}$ in exactly one edge and 7 vertices, and their remaining 2 vertices are distinct from those of any other copy of C_4^+.

We define our grow sequence according to Algorithm 3, where G is a minimal closed subgraph of $H^7(n,p)$.

Let C_1 be a copy of C_4^+ in G
$G_1 \leftarrow C_1$
$i \leftarrow 1$
while $G_i \neq G$ **do**
  $i \leftarrow i + 1$
  if *G_{i-1} contains an open edge* **then**
    $j \leftarrow$ smallest index $j < i$ such that C_j contains open edges
    $e \leftarrow$ the minimal open edge in C_j
    $C_i \leftarrow$ a copy of some hypergraph from $\mathcal{C}^*$ in G but not G_{i-1} which contains e as the attachment edge
  else
    if *there exists a copy C^* of a hypergraph in $\mathcal{C}^*$ in G, not contained in G_{i-1}, which intersects G_{i-1} in its attachment edge* **then**
      $C_i \leftarrow C^*$
    else
      $C_i \leftarrow$ some copy of C_4^+ not contained in G_{i-1} which intersects G_{i-1} in at least one edge
    end
  end
  $G_i \leftarrow G_{i-1} \cup C_i$
end

Algorithm 3: Decomposing minimal closed hypergraphs in the asymmetric case

Note that after each iteration of the while loop the only edges which are possibly open in G_i are those which are contained in a copy of C_4^+ but not one of $C_{6,20}$. For this reason if G_{i-1} contains open edges, then G must always contain a copy of $C_{6,20}$ which intersects each of these open edges. Further the remaining 19 edges of this copy of $C_{6,20}$ must also each intersect some copy of C_4^+. For this reason we can always find some hypergraph in $\mathcal{C}^*$ satisfying the requirements of Line 9. If G_i contains no open edges but is still not equal to G, then there must exist either a copy of $C_{6,20}$ (and thus of some $C^* \in \mathcal{C}^*$) or C_4^+ in G which is not contained in G_{i-1}, but is not edge-disjoint from it. The algorithm is therefore correct and terminates.

As in our first application of this grow sequence technique we need to establish some bounds for the contribution of regular and degenerate steps to our first moment argument (we define these precisely below). We begin with this lemma, which for degenerate steps essentially replicates (5.8), but in contrast to the symmetric case does not simply follow from an assumption of strict k-balancedness. We prove it in the next section.

Lemma 5.21. *For all $C^* \in \mathcal{C}^*$ we have that all $H \subsetneq C^*$ which contain the attachment edge of C^* satisfy*

$$\frac{e(C^*) - e(H)}{v(C^*) - v(H)} \geq \frac{76}{51} = m_7(C_4^+, C_{6,20}). \tag{5.12}$$

If the inequality is tight, then H consists of exactly the attachment edge of C^*. In this case it also holds that all 19 copies of C_4^+ in C^* intersect the copy of $C_{6,20}$ in exactly 7 vertices and 1 edge, and the remaining 2 vertices are distinct from those of any other copy of C_4^+.

As in Section 5.3 we distinguish various step types. C_1 is the first step. For a step C_i set $H_i := C_i \cap G_{i-1}$. If $(e(C_i) - e(H_i))/(v(C_i) - v(H_i)) = 76/51$ then we call the step regular, otherwise degenerate. We further call regular steps open or closed by whether they are chosen in line 9 or 12. The step chosen in line 14 is always degenerate.

Consider any open regular step C_i. By Lemma 5.21 its intersection H_i with G_{i-1} is exactly one edge, and thus we can bound the contribution of each such step by

$$|\mathcal{C}^*| v(C_i)^7 n^{v(C_i)-7} p^{e(C_i)-1} \leq |\mathcal{C}^*| 58^7 c^{76} n^{51-76\cdot 51/76} \leq c,$$

where c is the constant in Theorem 5.9 which we choose small enough for the above to hold, and small enough that $\log(c) < -1$ holds.

If C_i is regular closed we have

$$|\mathcal{C}^*|v(C_i)^7(v(C_i)\cdot i)^7 n^{v(C_i)-7}p^{e(C_i)-1} \le c\cdot(58\cdot i)^7.$$

Now consider the case of a degenerate step C_i which is a copy of some $C^* \in \mathcal{C}^*$. By Lemma 5.21 we can choose a constant $\alpha_1 > 0$ such that regardless of the choice of $C^* \in \mathcal{C}^*$ we have

$$v(C^*) - v(H_i) - \frac{e(C^*) - e(H_i)}{m_7(C_4^+, C_{6,20})} < -\alpha_1.$$

In the case of a degenerate step consisting of just a copy of C_4^+ (i.e. one as in Line 14) we can choose $\alpha_2 > 0$ such that for all $H_i \subsetneq C_4^+$, $e(H_i) \ge 1$, we have

$$v(C_4^+) - v(H_i) - \frac{e(C_4^+) - e(H_i)}{m_7(C_4^+, C_{6,20})} < -\alpha_2.$$

Note that this holds even for H_i being exactly one edge, as $m_7(C_4^+) > m_7(C_4^+, C_{6,20})$. We then set $\alpha = \min\{\alpha_1, \alpha_2\}$.

With this we obtain for any degenerate step C_i

$$|\mathcal{C}^*|v(C_i)^{v(H_i)}(i\cdot v(C_i))^{v(H_i)} n^{v(C_i)-v(H_i)}p^{e(C_i)-e(H_i)} \le i^{58}n^{-\alpha}.$$

As in the symmetric case we prove a bound on the length of a grow sequence in terms of the number of degenerate steps it contains. The proof of this lemma is also postponed to the next section.

Lemma 5.22. *Let S be a grow sequence of length ℓ.*

i) If S contains at most d degenerate steps, then $\ell < 61d + 1$.

ii) If a prefix S_i of S, $1 \le i \le \ell$, contains at most d degenerate steps, then S_i contains no closed regular steps F_j with $j > 61d + 1$.

We can now finish our first moment argument. Set $d_{\max} := 58/\alpha + 1$ and $L = 61\cdot d_{\max} + 1 = 3538/\alpha + 62$. By Lemma 5.22 all sequences longer than L must contain at least $d_{\max}$ degenerate steps. Set $\ell_{\max} := 58\log(n) + d_{\max} + 1$. We consider two cases: the first are those sequences which have their $d_{\max}$th degenerate step before $\ell_{\max}$. We truncate them after the $d_{\max}$th step. The second are those sequences whose $d_{\max}$th step appears after $\ell_{\max}$. We truncate these at length $\ell_{\max}$. Note that by Lemma 5.22 in both these cases closed regular steps can only happen in the first L steps of the sequence.

In the first case we obtain

$$\sum_{\ell=L+1}^{\ell_{\max}-1} n^{58} \binom{\ell}{d_{\max}} \left(\ell^{58} n^{-\alpha}\right)^{d_{\max}} \left((58 \cdot L)^7\right)^L = \\ = \mathcal{O}(\text{polylog}(n) \cdot n^{58} n^{-\alpha \cdot d_{\max}}) = o(1).$$

Here we bound the contribution of the first step by n^{58}, drop the contribution of $c < 1$ for all regular steps, and use the fact that only the first L steps may be closed regular.

In the second case we obtain

$$\sum_{d=0}^{d_{\max}} n^{58} \binom{\ell_{\max}}{d} \left((\ell_{\max})^{58} n^{-\alpha}\right)^d \left((58 \cdot L)^7\right)^L c^{\ell_{\max}-d} \\ = \mathcal{O}(\text{polylog}(n) \cdot n^{58} c^{\ell_{\max}-1-d}) = \mathcal{O}(\text{polylog}(n) \cdot n^{58(1+\log c)}) = o(1).$$

With this we proved that only grow sequences of length at most the constant L can appear in $H^7(n,p)$ for $p \leq cn^{m_7(C_4^+, C_{6,20})}$. For any such short sequence S it must hold that $m(H(S)) \leq m_7(C_4^+, C_{6,20})$, otherwise they are a.a.s. not contained in $H^7(n,p)$ by (5.5). The next lemma establishes that $C_4^+ \uplus C_{6,20}$ has a property equivalent to being Ramsey-density-obeying for the asymmetric Ramsey case and shows that we can always color such short sequences without a forbidden monochromatic copy of C_4^+ or $C_{6,20}$. This then concludes the proof of Theorem 5.9.

Lemma 5.23. *Let G be a 7-uniform hypergraph for which it holds that $m(G) \leq m_7(C_4^+, C_{6,20})$. Then $G \nrightarrow (C_4^+, C_{6,20})$.*

Proof. Without loss of generality we assume that we want to color G without a red C_4^+ or a blue $C_{6,20}$ (we call this a valid coloring). Recall that $m_7(C_4^+, C_{6,20}) = 76/51$, therefore by Lemma 5.16 G is at most 10-degenerate. It follows that we can choose an ordering $v_1, \ldots, v_{v(G)}$ of the vertices of G such that for $1 \leq i \leq v(G)$ we have that $\deg_{G_i}(v_i) \leq 10$, where G_i denotes the graph induced in G by $v_1, \ldots, v_i$.

We color G by induction on G_i, $1 \leq i \leq v(G)$. As G_1 contains no edges there exists a valid coloring of it. Consider G_i for some $2 \leq i \leq v(G)$. By induction there exists a valid coloring for all edges induced in G_i by G_{i-1}, and the only uncolored edges are those incident to v_i.

If v_i is not part of a copy of $C_{6,20}$ in G_i, then we simply color all edges incident to v_i blue and are done. Similarly, if v_i is contained in only one copy of $C_{6,20}$ in G_i, we color one of the edges contained in this copy of $C_{6,20}$ red and all other edges incident to v_i blue.

So assume that v_i is contained in $\ell \geq 2$ copies of $C_{6,20}$. Consider the subgraph H of $C_{6,20}$ induced by all edges incident to an arbitrary vertex of $C_{6,20}$. It holds that H contains 13 vertices and 7 edges, and that it contains 2 vertices each of degrees 1 to 6 and one vertex of degree 7. We call an edge in H an ℓ-edge if the minimum degree in H of the vertices it contains is ℓ. There are a pair each of 1, 2 and 3-edges and one single 4-edge. We call the 1-edges of H the *tail edges* or *tails*, and the 4-edge the *center edge*. Note that a tail and the center edge intersect in 4 vertices, while the two tails intersect in exactly 1.

Now consider the $\ell \geq 2$ copies of $C_{6,20}$ intersecting v_i in G_i. The subgraphs of these copies induced by the edges incident to v_i are copies of H and we denote them by $H_1, \ldots, H_\ell$.

Assume first that every H_i, $2 \leq i \leq \ell$, contains one of the tails or the center edge of H_1. Then we color both tails and the center edge of H_1 red and all other edges incident to v_i blue. This avoids blue copies of $C_{6,20}$ and does not create a red copy of C_4^+ either, as the minimum degree of C_4^+ is 2 and its edges have pairwise intersections of size 5 or more.

So assume that there exists some H_i, without loss of generality assume it is H_2, which contains none of the tails nor center edge of H_1. We claim that coloring the two 2-edges of H_1 red achieves our goal. Note that these two edges intersect in 3 vertices, so by the same argument as above these cannot create red copies of C_4^+. We claim that any other copy of H (and thus of $C_{6,20}$) must intersect at least one of these red edges.

To prove this first note that v_i is incident to at most 10 edges in G_i, 7 of which are contained in H_1. As by assumption H_2 must have 3 edges which are not contained in H_1 it follows that every edge incident to v_i must also be contained in H_1 or H_2.

Consider now the set of 4 tail edges from H_1 and H_2, and let H_i be some copy of H which contains $0 \leq j \leq 4$ of them.

The case $j = 0$ is not possible, as avoiding all 4 tail edges leaves only 6 edges which can be contained in H_i instead of the 7 edges that are required.

If $j = 1$ or $j = 2$ then H_i must contain all 6 resp. 5 of 6 non-tail edges. In both cases it contains at least one of the red edges.

If $j = 4$ then H_i contains two pairs of edges which intersect in exactly one vertex. However H contains only one such pair (the tail edges), so H_i cannot be a copy of H.

It remains to deal with the case $j = 3$. Assume that the set of chosen tail-edges contains both of those from H_1. These two edges must also be the tails of H_i, as they intersect in 1 vertex and the only pair of edges to do so in a copy of H are the tails. Further, if H_i avoids both red edges, it must contain every other edge of H_1 and additionally one tail and the center edge of H_2. Given one tail of a copy of H it holds that the 6 other edges intersect it in exactly $6, 5, \ldots, 1$ vertices. The intersections of one tail of H_1 with the edges also in H_1 cover the cases 5, 4, 3 and 1. This implies that the tail and center edge from H_2 must take over the roles of the missing red edges, which intersect a tail in 2 and 6 vertices respectively. However in a copy of H these two edges intersect in 3 and not 4 vertices, as is the case for the tail and center edge of H_2, which is a contradiction.

Assume now instead that H_i contains both tail edges from H_2 and only one of H_1. By the same argument as above we have that the two tails of H_2 are the tails of H_i. Assume that H_i does not contain the center edge of H_2. Then it must contain 4 non-tail edges from H_1. As there are only 3 non-red edges which are not tails in H_1 this is a contradiction. It remains to deal with the case that H_i contains both tails and the center edge c of H_2. This implies that c is also the center edge of H_i (by its intersections with the two tails). By assumption H_i avoids 1 tail and both 2-edges of H_1, it must therefore contain all other 4 edges of H_1. These are a tail edge t, both 3-edges e_3, e_3', and the center edge e_4 of H_1. The intersection of t with one of e_3 or e_3' (wlog: e_3) consists of 3 vertices. As both t and e_3 are not tails of H_i this implies that they must be its two 2-edges. This however makes e_3' the center edge of H_i, as it intersects both t and e_3 in 5 vertices, and thus $c = e_3'$. This is a contradiction as by assumption c is not contained in H_1. $\square$

5.7 Deterministic lemmas for the asymmetric case

In this section we prove lemmas 5.21 and 5.22.

We begin with the proof of Lemma 5.21.

Let C^* be any hypergraph in $\mathcal{C}^*$. Note that C^* contains exactly one copy of $C_{6,20}$ and several copies of C_4^+. As we use this fact many times in this proof we omit writing “a copy of” or “the copy of” in front of each mention of $C_{6,20}$ and C_4^+, and trade a little formal correctness for readability. We call the vertices of C^* from $C_{6,20}$ the *cycle vertices*, while we refer to all other vertices as *outside vertices*.

We first prove the following lemma, which deals with the case of H being just the attachment edge in Lemma 5.21.

Lemma 5.24. *Let $C^* \in \mathcal{C}^*$. Then*

$$\frac{e(C^*) - 1}{v(C^*) - 7} \geq \frac{76}{51}.$$

Further the inequality is tight if and only if each of the 19 copies of C_4^+ contains exactly 2 outside vertices and all of these are pairwise disjoint.

Proof. Let $C^* \in \mathcal{C}^*$ and arbitrarily label the 19 copies of C_4^+ that are attached to the 19 edges of $C_{6,20}$ as $C_{4,1}^+, \ldots, C_{4,19}^+$. Define

$$H_i := C_{4,i}^+ \cap \Big(C_{6,20} \cup \bigcup_{j=1}^{i-1} C_{4,j}^+ \Big).$$

(More precisely: H_i is given by performing the intersection for both the vertex and the edge sets.) Note that every H_i contains at least one edge, namely the one it shares with $C_{6,20}$. Set $e_i = e(H_i) - 1$ and $v_i = v(H_i) - 7$. Recall that $v(C_{6,20}) = e(C_{6,20}) = 20$ and $v(C_4^+) = 9$, $e(C_4^+) = 4$. With this we have

$$\begin{aligned}\frac{e(C^*) - 1}{v(C^*) - 7} &= \frac{e(C_{6,20}) - 1 + \sum_{i=1}^{19} \big(e(C_4^+) - e(H_i) \big)}{v(C_{6,20}) - 7 + \sum_{i=1}^{19} \big(v(C_4^+) - v(H_i) \big)} = \\ &= \frac{19 + \sum_{i=1}^{19} \big(4 - (e_i + 1) \big)}{13 + \sum_{i=1}^{19} \big(9 - (v_i + 7) \big)} = \frac{76 - \sum_{i=1}^{19} e_i}{51 - \sum_{i=1}^{19} v_i}. \end{aligned} \tag{5.13}$$

This completes the claim in the case that $v_i = e_i = 0$ for all i. Note that this corresponds to the case that the 19 copies of C_4^+ all add two new vertices and three new edges.

Note that by Proposition 2.1 it suffices to show that if v_i and e_i are both nonzero then $e_i / v_i < 76/51$. By construction, H_i is a subgraph of C_4^+.

Complete enumeration of all possible easily shows that whenever $v(H_i) > 7$ and $H_i \neq C_4^+$ we have

$$\frac{e(H_i) - 1}{v(H_i) - 7} = \frac{e_i}{v_i} \leq 1 < \frac{76}{51}. \tag{5.14}$$

This thus completes the proof of the lemma if none of the H_i is equal to C_4^+.

It remains to deal with the case that there exists some i, $1 \leq i \leq 19$, with $H_i = C_{4,i}^+$. In this case we say that H_i is *full*. Let $I := \{1 \leq i \leq 19 \mid H_i \text{ is full }\}$. We argue that if I is non-empty, then there exists some $j \in \{1, \dots, 19\} \setminus I$ such that

$$\frac{e_j + \sum_{i \in I} e_i}{v_j + \sum_{i \in I} v_i} < \frac{76}{51}, \tag{5.15}$$

which will complete the proof of the lemma.

To prove (5.15) fix $i \in I$ arbitrarily. Let $j < i$ be such that $C_{4,j}^+$ introduces the last edges and vertices that cause H_i to be full, i.e. we choose j as the maximum of $\{j_e \mid e \in E(H_i) \setminus E(C_{6,20})\}$, where $1 \leq j_e < i$ is the lowest index such that C_{4,j_e}^+ contains e.

First assume that H_j consists of just the edge of $C_{4,j}^+$ shared with $C_{6,20}$. In this case $C_{4,j}^+$ adds two new vertices and three new edges which are all not present in the preceding copies of C_4^+. This implies that *all* edges $e \in E(H_i) \setminus E(C_{6,20})$ have to be contained in H_j, which implies $C_{4,j}^+ = C_{4,i}^+$. However, as $C_{4,j}^+$ intersects $C_{6,20}$ in exactly one edge it must be the unique copy matched to this edge, and $C_{4,j}^+ \neq C_{4,i}^+$, which is a contradiction. Next assume that $v_i = 1$. Then H_j misses one vertex of $C_{4,j}^+$. As the minimum degree of C_4^+ is 2, this implies that H_j also misses at least 2 edges of $C_{4,j}^+$. Thus, $e_j \leq 1 = v_j$. Finally, assume that $v_i = 2$. As $C_{4,j}^+$ must introduce at least one new edge, we deduce also in this case that $e_j = e(H_j) - 1 \leq 2 = v_i$. As $|I| < 19$ one easily checks that

$$\frac{e_j + \sum_{i \in I} e_i}{v_j + \sum_{i \in I} v_i} = \frac{e_j + 3|I|}{v_j + 2|I|} \leq \frac{v_j + 3|I|}{v_j + 2|I|} \leq \frac{1 + 3|I|}{1 + 2|I|} < \frac{76}{51},$$

and we are done. □

As a second preliminary step we prove a lower bound on the number of edges that are incident to outside vertices.

Lemma 5.25. *Let V be a subset of the outside vertices of C^* and let E be the set of edges incident to V. It holds that $|E| \geq \frac{3}{2}|V|$.*

Proof. All edges of C^* are incident to either 0, 1 or 2 outside vertices, and accordingly we refer to them as 0-, 1- or 2-edges. Note that by construction each outside vertex has degree at least 2 and each is incident to at least one 1-edge. Let $A \subseteq V$ be the subset of those vertices in V which are incident to at least two 1-edges, and let $B = V \setminus A$ be the remaining vertices. It follows that E contains at least $2|A| + |B|$ many 1-edges. The number of 2-edges incident to B is at least $|B|/2$, as each edge is incident to at most 2 vertices from B. In total we obtain

$$|E| \geq 2|A| + |B| + \frac{|B|}{2} \geq \frac{3}{2}|V|,$$

regardless of how V splits into A and B. □

With these tools at hand we can now prove Lemma 5.21.

Proof of Lemma 5.21. Lemma 5.24 covered the case in which H consists only of the attachment edge. For the remainder of the proof we thus assume that $e(H) > 1$. We also assume that H is such that it minimizes $(e(C^*) - e(H))/(v(C^*) - v(H))$.

If $H \subsetneq C^*$ contains the entire copy of $C_{6,20}$, then $V(C^*) \setminus V(H)$ consists only of outside vertices and Lemma 5.25 thus implies that

$$\frac{e(C^*) - e(H)}{v(C^*) - v(H)} \geq 3/2.$$

If this is not the case then H avoids $\ell \geq 1$ vertices of $C_{6,20}$. As H contains the attachment edge, we know that $\ell \leq 13$. It is easy to see that any subset of size $\ell \leq 13$ of the vertices of $C_{6,20}$ intersects at least $\ell + 6$ edges of $C_{6,20}$. Consider now the hypergraph H' that arises by adding the ℓ missing vertices from the $C_{6,20}$ and the edges incident to them. By the minimality of H we have

$$\frac{e(C^*) - e(H)}{v(C^*) - v(H)} \leq \frac{e(C^*) - e(H')}{v(C^*) - v(H')} = \frac{(e(C^*) - e(H)) - (e(H') - e(H))}{(v(C^*) - v(H)) - \ell}. \tag{5.16}$$

As $e(H') - e(H) \geq \ell + 6$ and $(\ell + 6)/\ell \geq 3/2$ for all $1 \leq \ell \leq 12$, it follows that (5.16) can only hold if $(e(C^*) - e(H))/(v(C^*) - v(H) \geq 3/2$. This proves the claim for all $\ell \leq 12$.

It remains to deal with the case of H avoiding 13 vertices of $C_{6,20}$. In this situation H contains, besides the attachment edge, only outside vertices of C^* and edges which contain only outside vertices and vertices from the attachment edge. Note that in particular H contains no copy of C_4^+, as such a copy would intersect $C_{6,20}$ only on the attachment edge, which is not possible by the definition of $\mathcal{C}^*$.

Set $V' := V(H) \cup V(C_{6,20})$ and let E' be the set of edges from C^* induced by V'. Let H' be the hypergraph with vertex and edge set V' resp. E'. Note that $e(H') \geq e(H) + 19$ and $v(H') = v(H) + 13$. Set $e' := e(H') - e(H) - 19$.

By the minimality of H it holds that

$$\frac{e(C^*) - e(H)}{v(C^*) - v(H)} \leq \frac{e(C^*) - e(H')}{v(C^*) - v(H')} = \frac{e(C^*) - e(H) - (19 + e')}{v(C^*) - v(H) - 13}.$$

If $e' > 0$, then $(19 + e')/13 > 3/2$, and we are done by the same argument as above.

It remains to deal with the case $e' = 0$. Note that this case is only possible if H avoids at least one outside vertex of every copy of C_4^+, as a full set of outside vertices together with the vertices of $C_{6,20}$ would induce the corresponding copy of C_4^+ in the construction of H' and imply $e' > 0$. Let v be an arbitrary outside vertex of H. Let C be any copy of C_4^+ in C^* containing v. By the above C must contain a second outside vertex not contained in H. Let H'' be the hypergraph obtained by adding v to H' together with the 2 edges incident to it in C. It holds that $e(H'') = e(H) + 21$ and $v(H'') = e(H) + 14$. As $21/14 = 3/2$ we are again done by the same argument as above. □

We now prove Lemma 5.22. The proof is very similar to the symmetric case, so we provide only a sketch highlighting the differences to the symmetric case proof in Section 5.4.

Let S be any grow sequence of length ℓ, and S_i, $i \leq \ell$ be a prefix of it. We call a step F_j, $j \leq i$, *fully open* in S_i if it is a regular step which contains $19 \cdot 3 = 57$ open edges (i.e. every edge of a C_4^+ not intersecting $C_{6,20}$) and no other step $F_{j'}$, $j < j' \leq i$, intersects any vertex of F_j which is not in the attachment edge of F_j. Note that 57 is the maximum number of edges which can be open in any regular step.

By Lemma 5.12 attaching a regular step F_i to G_{i-1} creates exactly one new copy of $C_{6,20}$, namely the one contained in F_i itself. To see this note that $C_{6,20}$ is strictly 7-balanced, and that $\gamma = 6$. Any other copy of $C_{6,20}$ would thus have to have an edge intersecting the attachment edge of F_i in 7 vertices, while still being distinct from it. This is a contradiction. By Lemma 5.21 all vertices of a regular step F_j which are not contained in the attachment edge have degree 2, and can therefore not be part of some additional $C_{6,20}$.

This implies that attaching a regular step F_i to G_{i-1} can close at most one open edge: the one corresponding to the attachment edge of F_i. With this it is straightforward to check that an equivalent lemma as Lemma 5.14 holds with the bound of $e(F) - 1$ steps replaced by one of 57 consecutive regular steps. The reason is that a fully open step has 57 open edges, and every subsequent regular step closes one of them.

Similarly Lemma 5.15 holds in our case with the bound in (5.10) replaced by

$$\mathrm{fo}(S_i) \geq \mathrm{reg}(S_i)\cdot\left(1-\frac{1}{57}\right)-\deg(S_i)\cdot v(C^*) \geq \mathrm{reg}(S_i)\cdot\left(1-\frac{1}{57}\right)-\deg(S_i)\cdot 58. \tag{5.17}$$

We obtain this by replacing the bound of $e(F) - 1$ which comes from Lemma 5.14 by the one of 57 we established above, and replace $v(F)$ by the maximum number of vertices in any $C^* \in \mathcal{C}^*$, which is 58. The proof of the inequality is identical.

Let S be a sequence of length ℓ with at most d degenerate steps. Since S reconstructs a closed hypergraph, it contains no open edges and in particular no fully open steps. It follows from (5.17) that

$$58 \cdot \deg(S) \geq \mathrm{reg}(S) \cdot \left(1 - \frac{1}{57}\right).$$

It holds that $\mathrm{reg}(S) = \ell - d - 1$, and thus

$$58d \geq (\ell - d - 1) \cdot \frac{56}{57}.$$

Solving for ℓ proves Lemma 5.22.

5.8 Open problems

Our definition of Ramsey-density-obeying does not make it particularly easy to check whether a given graph F satisfies it or not. While in principle it is "just" a finite enumeration problem, it may still be too large to carry out. Ideally one would wish for a more effective way of proving that a given hypergraph F is Ramsey-density-obeying, such as e.g. in the case of graphs, where a finite list of not Ramsey-density-obeying classes of graphs exists. We can prove [Tho13] that e.g. hypergraphs with high minimum degree or high chromatic number are Ramsey-density-obeying, but so far we have not found a complete characterization. In this context it is interesting to note that every example of a not Ramsey-density-obeying hypergraph we are aware of is either a path of length 3 (or its generalization to hypergraphs) or "star-like". By this we mean that there always is at least 1 vertex which is shared by every edge. It may be that every strictly k-balanced hypergraph with no vertex common to all edges is Ramsey-density-obeying, but we do not have a proof for this. We also do not know if the converse holds in some form.

Another open question is whether it is possible to have other different "types" of thresholds for the symmetric Ramsey problem in hypergraphs. In the graph case the possible thresholds are either given by not Ramsey-density-obeying graphs (those with a small counterexample) or are "as expected". We have proved in Theorem 5.9 that for hypergraphs there is at least one additional type of threshold which is not determined by a small and local counterexample, but which is also not "as expected" (i.e. our asymmetric example).

As far as we know Theorem 5.9 is the first non-trivial example of a threshold for the asymmetric Ramsey game in hypergraphs. Except for Lemma 5.21 the proof of it could fairly easily be generalized to at least a much broader class of hypergraphs. However it seems that even for the graph case such a statement is difficult to generalize.

Chapter 6

Balanced coloring games in random graphs

In this chapter we examine the balanced Ramsey game in both its edge-coloring and vertex-coloring version. We contrast it to the Achlioptas game and find that for the vertex case they have the same threshold, while for the edge case the thresholds differ. This is joint work with Reto Spöhel. An extended abstract has appeared in [GS11], and the full version in [GS14].

6.1 Introduction

6.1.1 The balanced Ramsey game

Recall that in the balanced Ramsey game we start with the empty graph on n vertices, then in each step r new edges are sampled uniformly at random from all non-edges and inserted into the graph. The player – we call her Painter – has r colors at her disposal and must color these r edges immediately subject to the restriction that each color is assigned to exactly one of the r edges. Her goal is to avoid creating a monochromatic copy of some fixed graph F for as long as possible.

The typical duration of this game when played with an optimal strategy is formalized by the notion of its threshold function $N_0(F, r, n)$. Specifically, we say that $N_0(F, r, n)$ is a threshold function for the game (for a fixed graph F and a fixed integer $r \geq 2$) if for any function $N(n) \ll N_0$, Painter can a.a.s. 'survive' for at least N steps using an appropriate strategy, and if for any $N(n) \gg N_0$, Painter a.a.s. cannot survive for more than N steps regardless of her strategy. Note that this defines the threshold function only up to constant factors; therefore, whenever we compare two threshold functions and e.g. say that one is strictly higher than the other this refers to their orders of magnitude.

Standard arguments show that such a threshold function always exists for games of this type (see [MSS09a, Lemma 2.1]). Therefore the goal when studying these games usually is to determine their threshold function *explicitly*. In [MMS07], Marciniszyn et al. determined the threshold function of the balanced Ramsey game for the case when F is a cycle of arbitrary fixed length, and $r = 2$ colors are available. For example, the threshold of the balanced Ramsey game when $F = C_3$ is a triangle and $r = 2$ was shown to be $N_0(C_3, 2, n) = n^{6/5}$. More recently, Prakash et al. [PST09] extended these results to an arbitrary number of colors $r \geq 2$. In particular, their work yields the first threshold results for the case where $F = K_\ell$ is a complete graph of size at least 4 (and r is large enough; specifically, their result requires $r \geq \ell$).

6.1.2 The Achlioptas game

In the Achlioptas game we start with an empty graph on n vertices. In each step, r edges chosen uniformly at random from all edges never seen before are revealed. The player has to choose exactly one of these edges for inclusion in the graph; the remaining $r-1$ edges are discarded. The player's goal is to avoid creating a copy of some fixed graph F for as long as possible. Note that this can be seen as a balanced Ramsey game with relaxed rules such that the player only needs to worry about copies of F in the first color and can ignore the other $r-1$ colors. As an immediate consequence, for any F and r the threshold of the Achlioptas game is an upper bound on the threshold of the balanced Ramsey game.

This game was first studied by Krivelevich et al. in [KLS09]. Mütze et al. [MST11] recently determined the general threshold function of this game, valid for any fixed graph F and any fixed integer $r \geq 2$. The general threshold formula turns out to be considerably more complicated than the preliminary results of [KLS09] suggest.

It follows from known results that if F is e.g. a star or a path, the balanced Ramsey and the Achlioptas game have different thresholds (see Section 6.5 for an example). However, for all non-forests F where both thresholds are known (i.e. for all cases covered by Prakash et al. [PST09]), the two thresholds coincide, and so far it was unknown whether in fact the two thresholds coincide for any non-forest F and any $r \geq 2$. This question was raised explicitly in Krivelevich et al. [KSS10]. We answer this question negatively in this work.

Theorem 6.1. *There is an infinite family of non-forests F for which, for any fixed integer $r \geq 2$, the balanced Ramsey game has a strictly lower threshold than the Achlioptas game.*

The simplest non-forest graph F for which we show that the two online thresholds differ consists of three triangles joined at a common vertex, cfr. Figure 6.2(a) on page 101.

Theorem 6.1 is in contrast with known results on the *offline* problems corresponding to the two online games discussed here: As shown in [KSS10], the two offline problems have the same threshold for 'almost all' non-forests F, in particular for 'most' graphs of the infinite family from Theorem 6.1.[1]

[1] The result is proven for all non-forests F that have a strictly 2-balanced subgraph $H \neq K_3$.

6.1.3 Vertex analogues

Both the balanced Ramsey and the Achlioptas game have a natural vertex analogue, where the player is presented with r new *vertices* (instead of edges) in each step. At the start of these games, a random graph $G(n,p)$ on vertex set $\{v_1,\ldots,v_n\}$ is generated, hidden from the player's view, by including each of the $\binom{n}{2}$ possible edges with some fixed probability $p = p(n)$ independently. We assume that r divides n. In each step of the game, the r next consecutive vertices are revealed, along with all edges induced by the vertices revealed so far. Thus after i steps, the player sees exactly the random edges induced by $v_1,\ldots,v_{ir}$.

In the balanced Ramsey game the player has to assign each of r available colors to exactly one of the r new vertices at each step, without completing a (vertex-)monochromatic copy of some fixed graph F. In the Achlioptas game, she has to select one of the r new vertices, and the $r-1$ remaining vertices are discarded along with all incident edges. Again the player's goal is to avoid creating a copy of some fixed graph F.

In both cases we are interested in finding explicit threshold functions $p_0 = p_0(F,r,n)$ such that (i) for any function $p(n) \ll p_0$ there is a strategy which a.a.s. allows the player to color (resp. choose from) all n vertices without creating a (monochromatic) copy of F, and (ii) for any $p(n) \gg p_0$ every possible player strategy a.a.s. fails to do so. (The mere existence of such threshold functions can again be shown similarly to [MSS09a, Lemma 2.1].)

Prakash et al. [PST09] proved results analogous to those discussed above for the edge-coloring setting also for the vertex case. Moreover, also the results of Mütze et al. [MST11] for the Achlioptas game translate with minimal changes to the vertex setting, even though this is not made explicit in their work. (We will elaborate on this in Section 6.2 below and in Section 6.6 at the end of this chapter.) To sum up, in the literature the vertex and the edge case of the two games are equally well understood, and the known results for them are in complete analogy to each other.

As we shall see, this pattern breaks down in the general case: We prove that in the vertex case the thresholds of the balanced Ramsey and the Achlioptas game coincide for all graphs F and all $r \geq 2$. This is in contrast with our result for the edge case given in Theorem 6.1.

Theorem 6.2 (Main result)**.** *For all graphs F and all $r \geq 2$, the vertex versions of the balanced Ramsey game and the Achlioptas game have the same threshold.*

We give the explicit threshold formula of the two games in Section 6.2 below.

6.1.4 Organization of this chapter

Recall that the threshold of the (vertex) Achlioptas is always an upper bound on the threshold of the (vertex) balanced Ramsey game. Hence to prove Theorem 6.2 it suffices to give an upper bound on the threshold of the vertex Achlioptas game and a matching lower bound on the vertex balanced Ramsey game.

In Section 6.2 we outline how the results of Mütze et al. [MST11] on the edge Achlioptas game, including their upper bound proof, translate to the vertex setting. The complete proofs for these results are given for reference in the last section, as they follow their edge counterparts quite closely and are not the main contribution of this work. In Section 6.3, we adapt some key concepts from [MST11] to the vertex setting. In Section 6.4, we then use these to prove the desired matching lower bound for the vertex balanced Ramsey game. Finally, we prove Theorem 6.1 concerning the edge case in Section 6.5.

6.2 On the vertex Achlioptas game

In this section we adapt the formalism and the results of Mütze et al. [MST11] from the edge to the vertex case. The proofs are very similar; and we reproduce them in the last section. We also refer the reader to [MST11] for a more in-depth discussion of the intuition behind our threshold formulas.

A (vertex-)*ordered graph* is a pair (H, π), where H is a graph, $h := v(H)$, and $\pi : V(H) \to \{1, \ldots, h\}$ is an ordering of the vertices of H, conveniently denoted by its preimages, $\pi = (\pi^{-1}(1), \ldots, \pi^{-1}(h))$. In the context of the vertex Achlioptas or balanced Ramsey game, we interpret the ordering $\pi =: (u_1, \ldots, u_h)$ as the order in which the vertices of H appeared in the process, where u_h is the vertex that appeared first (the "oldest" vertex) and u_1 is the vertex that appeared last (the "youngest" vertex). We denote by $\Pi(V(H))$ the set of all possible orderings of the vertices of H, and by

$$\mathcal{S}(F) := \big\{(H, \pi) \mid H \subseteq F \land \pi \in \Pi(V(H))\big\}$$

the set of all ordered subgraphs of F. For some ordered graph (H,π) and a subgraph $J \subseteq H$, we denote by $\pi|_J$ the order on the vertices of J induced by π. Given an ordered graph (H,π), $\pi = (u_1,\ldots,u_h)$, we denote by $H \setminus \{u_1,\ldots,u_i\}$ the graph obtained from H by removing the vertices $u_1,\ldots,u_i$ and all edges that contain at least one one of these vertices. (In other words, $H \setminus \{u_1,\ldots,u_i\}$ is the subgraph of H induced by the vertices $u_{i+1},\ldots,u_h$.) We use $u \in H$ as a shorthand notation for $u \in V(H)$.

For any graph H, we use the notations $e(H) := |E(H)|$ and $v(H) := |V(H)|$. For any nonempty ordered graph (H_1,π), $\pi = (u_1,u_2,\ldots,u_h)$, any sequence of subgraphs $H_2,\ldots,H_h \subseteq H_1$ with $H_i \subseteq H_1 \setminus \{u_1,\ldots,u_{i-1}\}$ and $u_i \in H_i$ for all $2 \le i \le h$, and any integer $r \ge 2$ define coefficients $c_i = c_i((H_1,\pi),H_2,\ldots,H_h,r)$ recursively by

$$\begin{aligned} c_1 &:= r, \\ c_i &:= (r-1)\cdot\sum_{j=1}^{i-1} c_j \mathbf{1}_{\{u_i \in H_j\}}, \qquad 2 \le i \le h \end{aligned} \tag{6.1}$$

(where $\mathbf{1}_{\{u_i \in H_j\}} = 1$ if $u_i \in H_j$ and $\mathbf{1}_{\{u_i \in H_j\}} = 0$ otherwise), and set

$$d^{r*}(H_1,\pi) := \max_{\substack{H_2,\ldots,H_h \\ \forall i \ge 2: H_i \subseteq H_1 \setminus \{u_1,\ldots,u_{i-1}\} \wedge u_i \in H_i}} \frac{\sum_{i=1}^h c_i e(H_i)}{1 + \sum_{i=1}^h c_i\big(v(H_i)-1\big)}. \tag{6.2}$$

Furthermore, we set for any integer $r \ge 2$ and any nonempty graph F,

$$m^{r*}(F) := \min_{\pi \in \Pi(V(F))} \max_{H_1 \subseteq F} d^{r*}(H_1, \pi|_{H_1}). \tag{6.3}$$

With these notations and definitions, the main result of [MST11] translates to the following statement for the vertex Achlioptas case:

Theorem 6.3. *Let F be a fixed nonempty graph, and let $r \ge 2$ be a fixed integer. Then the threshold of the vertex Achlioptas game with parameters F and r is*

$$p_0(F,r,n) = n^{-1/m^{r*}(F)}.$$

In particular, if $p(n) \gg n^{-1/m^{r*}(F)}$, the player a.a.s. loses the vertex Achlioptas game with parameters F and r, regardless of her strategy.

As discussed in the introduction, this result also yields an upper bound of $n^{-1/m^{r*}(F)}$ on the threshold of the vertex balanced Ramsey game with parameters F and r. We will prove a matching lower bound in the next section. We now present an alternative formulation of Theorem 6.3 that is more convenient for this lower bound proof. Again we refer to [MST11] for a discussion of the advantages of this alternative viewpoint.

Given an ordered graph (H, π), $\pi = (u_1, \ldots, u_h)$, we use $H \setminus u_1$ as a shorthand notation for $H \setminus \{u_1\}$, and $\pi \setminus u_1$ as a shorthand notation for $\pi|_{H\setminus\{u_1\}}$. As usual we denote for $u \in V(H)$ by $\deg_H(u)$ the degree of u in H.

For a fixed integer $r \geq 2$ and a fixed real value $0 \leq \theta \leq 2$ we recursively define for any ordered graph (H, π), $\pi = (u_1, \ldots, u_h)$, the following quantity:

$$\lambda_{r,\theta}(H,\pi) := \begin{cases} 0, & \text{if } v(H) = 0 \\ 1 + \Big(\lambda_{r,\theta}(H \setminus u_1, \pi \setminus u_1) - \theta \cdot \deg_H(u_1)\Big) & \\ \quad + (r-1) \cdot \min\limits_{\substack{J \subseteq H \\ u_1 \in J}} \Big(\lambda_{r,\theta}(J \setminus u_1, \pi|_{J\setminus u_1}) - \theta \cdot \deg_J(u_1)\Big), & \\ & \text{otherwise.} \end{cases} \tag{6.4}$$

We further define for r and θ as before and any F the quantity

$$\Lambda_{r,\theta}(F) := \max_{\pi \in \Pi(V(F))} \min_{H \subseteq F} \lambda_{r,\theta}(H, \pi|_H). \tag{6.5}$$

It is straightforward to check that as a function of θ for a fixed r and a fixed nonempty graph (H, π) respectively F, both $\lambda_{r,\theta}(H, \pi)$ and $\Lambda_{r,\theta}(F)$ are continuous, piecewise linear with integer coefficients, and non-increasing. Furthermore, both functions have a unique rational root.

Analogously to [MST11] one can prove:

Theorem 6.4. *Let F be a fixed nonempty graph, and let $r \geq 2$ be a fixed integer. Let $\theta^* = \theta^*(F, r)$ be the unique solution of*

$$\Lambda_{r,\theta}(F) \stackrel{!}{=} 0,$$

where $\Lambda_{r,\theta}(F)$ is defined in (6.4) and (6.5). Then we have

$$m^{r*}(F) = \frac{1}{\theta^*(F, r)} .$$

Consequently, the threshold of the vertex Achlioptas game with parameters F and r can be written as

$$p_0(F, r, n) = n^{-\theta^*(F,r)}.$$

6.3 r-matched graphs

In this section we adapt some key notions concerning r-(edge-)matched graphs introduced in [KSS10] and [MST11] to the vertex setting studied here. We will need these concepts in our proof of a lower bound on the vertex balanced Ramsey threshold.

Definition 6.5 (r-matched graph). *An* r-(vertex-)matched graph $G = (V, E, \mathcal{K})$ *is a (simple, undirected) graph with vertex set V and edge set E together with a partition $\mathcal{K}$ of V into sets of size r, the* r-sets. *With $\kappa(G) := |\mathcal{K}| = |V|/r$ we denote the number of r-sets of G. We refer to the (non-r-matched) graph $G' = (V, E)$ as the* underlying graph of G.

We extend standard notions like graph isomorphism, subgraph containment etc. to r-matched graphs in the obvious way.

Recall that the vertex Achlioptas and vertex balanced Ramsey game is played on a binomial random graph $G(n, p)$ on vertex set $\{v_1, \ldots, v_n\}$ that is initially hidden from the player's view and revealed r vertices at a time. We denote by G_i the graph induced by $\{v_1, \ldots, v_{ir}\}$ (i.e. the graph visible to the player after i steps), viewed as an (uncolored) r-matched graph with partition $\mathcal{K} = \{\{v_1, \ldots, v_r\}, \{v_{r+1}, \ldots, v_{2r}\}, \ldots, \{v_{(i-1)r+1}, \ldots, v_{ir}\}\}$. In particular, $G_{n/r}$ is the random graph $G(n, p)$ generated before the game starts, *viewed as an r-matched graph* with the obvious partition. We denote a generic instance of such a random r-matched graph by $G^r(n, p)$ in the following.

In our lower bound proof we will need the following simple lemma.

Lemma 6.6. *Let $r \geq 2$ be a fixed integer, and let F be a fixed r-matched graph with at least one edge. Then the expected number of copies of F in $G^r(n, p)$ is $\Theta(n^{\kappa(F)} p^{e(F)})$.*

Proof. There are $\binom{n/r}{\kappa(F)} \cdot \Theta(1) = \Theta(n^{\kappa(F)})$ possible occurrences of F in $G^r(n, p)$, and each of them appears with probability $p^{e(F)}$. □

For $r \geq 2$, any r-matched graph F and $0 \leq \theta \leq 2$ let

$$\mu_{r,\theta}(F) := \kappa(F) - \theta \cdot e(F). \tag{6.6}$$

Note that, by the above lemma, for $p := n^{-\theta}$ the expected number of copies of F in $G^r(n, p)$ is of order $n^{\mu_{r,\theta}(F)}$.

6.4 A matching lower bound on the vertex balanced Ramsey threshold

In this section we prove the main contribution of this work, a lower bound on the threshold of the balanced Ramsey game that matches the upper bound given by Theorem 6.3. In view of Theorem 6.4, it suffices to prove the following statement.

Theorem 6.7. *Let F be a fixed nonempty graph, and let $r \geq 2$ be a fixed integer. Let $\theta^* = \theta^*(F,r)$ be the unique solution of*

$$\Lambda_{r,\theta}(F) \stackrel{!}{=} 0, \tag{6.7}$$

where $\Lambda_{r,\theta}(F)$ is defined in (6.4) *and* (6.5)*. Then for all $p \ll n^{-\theta^*}$ there exists a strategy such that Painter can a.a.s. win the vertex balanced Ramsey game with parameters F and r.*

We now describe the general coloring strategy for which we will prove Theorem 6.7. The strategy is a natural extension of the one proposed in [MST11] for the (edge or vertex) Achlioptas game; and we use very similar notations and conventions in the following. Note however that the analogous extension of the *edge* Achlioptas strategy fails to yield a similar lower bound, cfr. Theorem 6.1 and its proof in Section 6.5.

Crucially, our strategy keeps track of the *order* in which copies of subgraphs appear on the board. We say that the board contains a (monochromatic) copy of (H,π), $\pi = (u_1,\ldots,u_h)$, if it contains a (monochromatic) subgraph isomorphic to H whose vertices appeared in the order specified by π (with u_h being the first and u_1 being the last vertex to appear).

Let $r \geq 2$ and $0 \leq \theta \leq 2$ be arbitrary but fixed. (Eventually we will set $\theta = \theta^*(F,r)$, but for the moment it is more convenient to work with an arbitrary θ.) We denote with C the set of available colors. Consider a fixed step of the game, and let R denote the r-set presented to the player in that step. (We have $R = \{v_{(i-1)r+1},\ldots,v_{ir}\}$ for some i, $1 \leq i \leq n/r$.) Painter's decision in this step can be formalized as choosing a perfect matching in the complete bipartite graph B with parts C and R, where each edge corresponds to assigning a color to a vertex. We say that a perfect matching M *closes a copy* of some graph $(H,\pi) \in \mathcal{S}(F)$ if coloring R according to M creates a monochromatic copy of (H,π) on the board (clearly, then the last vertex of H according to π is in R).

Painter's strategy now is the following: She partitions B into r disjoint perfect matchings $M_1, \ldots, M_r$ arbitrarily. (By an easy application of the marriage theorem, this is always possible.) For each of these matchings she determines the value

$$d(M) := \min\bigl\{\lambda_{r,\theta}(H,\pi) \mid \ (H,\pi) \in \mathcal{S}(F) \wedge M \text{ closes a copy of } (H,\pi)\bigr\}, \tag{6.8}$$

and chooses the matching for which this value is maximal.

If there is not a unique maximum, ties are broken according to the following somewhat technical criterion. Consider the directed graph $\mathcal{G} = \mathcal{G}(F)$ with vertex set $\mathcal{S}(F)$ and arcs given by proper (ordered) subgraph inclusion; i.e., from every vertex (H,π) there are arcs to all vertices $(J,\pi|_J)$ with $J \subsetneq H$. Clearly, $\mathcal{G}$ contains no directed cycles. We extend $\mathcal{G}$ to a graph $\mathcal{G}' = \mathcal{G}'(F,r,\theta)$ by first connecting every pair of distinct vertices (H_1,π_1), (H_2,π_2) for which $\lambda_{r,\theta}(H_1,\pi_1) = \lambda_{r,\theta}(H_2,\pi_2)$ with an (undirected) edge, and then orienting these additional edges in such a way that the directed graph $\mathcal{G}'$ remains acyclic. (It is easy to see that this is always possible.) Note that for every fixed $\lambda_0 \in \mathbb{R}$ this yields a total ordering on all graphs (H,π) with $\lambda_{r,\theta}(H,\pi) = \lambda_0$. We say that (H_1,π_1) is higher than (H_2,π_2) in this ordering if the corresponding arc in $\mathcal{G}'$ is directed from (H_1,π_1) to (H_2,π_2).

Our strategy breaks ties according to this ordering: Whenever we have a choice between different perfect matchings with the same value $d(M)$, then for each such matching we consider the set of ordered graphs

$$\mathcal{J}(M) := \arg\min\{\lambda_{r,\theta}(H,\pi) \mid (H,\pi) \in \mathcal{S}(F) \wedge M \text{ closes a copy of } (H,\pi)\} \tag{6.9}$$

and, among these, we let $J(M) \in \mathcal{J}(M)$ denote the graph that is *lowest* in the total ordering for $\lambda_0 := d(M)$. Then we select the matching M for which $J(M)$ is *highest* in the total ordering for λ_0.

The next lemma states a witness graph invariant that is crucial in our proof of Theorem 6.7. Note that the statement of the lemma is purely deterministic.

Lemma 6.8. *Let $r \geq 2$ be an integer and $0 \leq \theta \leq 2$ be fixed. Following the above vertex coloring strategy ensures that the following invariant is maintained throughout the game for some $v_{\max} = v_{\max}(F,r,\theta)$:*

The graph G_i contains a copy of some r-matched graph K' with $v(K') \leq v_{\max}$ and

$$\mu_{r,\theta}(K') < 0,$$

or for every $(H, \pi) \in \mathcal{S}(F)$ we have that each monochromatic copy of (H, π) on the board is contained in an r-matched subgraph H' of G_i with $v(H') \le v_{\max}$ and

$$\mu_{r,\theta}(H') \le \lambda_{r,\theta}(H, \pi), \tag{6.10}$$

where $\mu_{r,\theta}()$ and $\lambda_{r,\theta}()$ are defined in (6.6) *and* (6.4)*, respectively.*

We postpone the proof of Lemma 6.8 and show first how it implies Theorem 6.7.

Proof of Theorem 6.7*.* Let $\theta^* = \theta^*(F, r)$ be defined as in the theorem. We show that the above strategy for $\theta := \theta^*$ allows Painter to win a.a.s. for all $p \ll p_0(r, F, n) = n^{-\theta^*}$.

By the definition of θ^* (cf. (6.5) and (6.7)) we have that for each possible ordering π of the vertices of F there exists some pair $(H, \pi|_H) \in \mathcal{S}(F)$ such that $\lambda_{r,\theta^*}(H, \pi|_H) \le 0$. According to Lemma 6.8 the following holds for each such $(H, \pi|_H)$: If the final board contains a monochromatic copy of $(H, \pi|_H)$, then $G_{n/r}$ contains an r-matched graph K' of size at most $v_{\max}$ and $\mu_{r,\theta^*}(K') < 0$, or an r-matched graph H', again of size at most $v_{\max}$, satisfying

$$\mu_{r,\theta^*}(H') \le \lambda_{r,\theta^*}(H, \pi|_H) \le 0 \ .$$

This yields a family $\mathcal{W} = \mathcal{W}(F, \pi, r)$ of r-matched graphs W' satisfying $\mu(W') \le 0$ and $v(W') \le v_{\max}$ such that, deterministically, $G_{n/r}$ contains a graph from $\mathcal{W}$ if the final board contains a monochromatic copy of (F, π) (and hence also a copy of $(H, \pi|_H)$). It follows that $G_{n/r}$ contains a graph from $\mathcal{W}^* = \mathcal{W}^*(F, r) := \cup_{\pi \in \Pi(E(F))} \mathcal{W}(F, \pi, r)$ if the final board contains a monochromatic copy of F. Moreover, as no graph in $\mathcal{W}^*$ has more than $v_{\max}$ vertices, the size of $\mathcal{W}^*$ is bounded by a constant depending only on F and r. As $G_{n/r}$ is distributed as a random r-matched graph $G^r(n, p)$, we obtain with Lemma 6.6, the definition of $\mu_{r,\theta^*}()$ in (6.6), and the fact that $\mu_{r,\theta^*}(W') \le 0$ for all $W' \in \mathcal{W}^*$, that for $p \ll n^{-\theta^*}$ the expected number of copies of graphs from $\mathcal{W}^*$ in $G_{n/r}$ is of order

$$\sum_{W' \in \mathcal{W}^*} n^{\kappa(W')} p^{e(W')} \ll \sum_{W' \in \mathcal{W}^*} n^{\mu_{r,\theta}(W')} \le |\mathcal{W}^*| \cdot n^0 = \Theta(1).$$

It follows from Markov's inequality that a.a.s. $G_{n/r} \cong G^r(n, p)$ contains no r-matched graph from $\mathcal{W}^*$. Consequently a.a.s. the final board contains no monochromatic copy of F. $\square$

For the proof of Lemma 6.8 we require the following technical lemma concerning the minimization in the definition of $\lambda_{r,\theta}()$. The proof is straightforward and analogous to [MST11, Lemma 10].

Lemma 6.9. *Let $r \geq 2$ be an integer, $0 \leq \theta \leq 2$ fixed, and let $\mathcal{F}$ be a family of ordered graphs with the property that if some (H, π), $\pi = (u_1, \dots, u_h)$, is in $\mathcal{F}$ then for every subgraph $J \subseteq H$ with $u_1 \in J$ also $(J, \pi|_J)$ is in $\mathcal{F}$. Then for $\lambda_{r,\theta}()$ as defined in (6.4) we have*

$$\operatorname*{arg\,min}_{(H,\pi)\in\mathcal{F}} \lambda_{r,\theta}(H,\pi) = \operatorname*{arg\,min}_{(H,\pi)\in\mathcal{F}} \Big(\lambda_{r,\theta}(H \setminus u_1, \pi \setminus u_1) - \theta \cdot \deg_H(u_1)\Big) \subseteq \mathcal{F}, \tag{6.11}$$

and all ordered graphs $(\hat{J}, \hat{\pi})$, $\hat{\pi} = (\hat{u}_1, \dots, \hat{u}_j)$, in the family (6.11) *satisfy*

$$\lambda_{r,\theta}(\hat{J}, \hat{\pi}) = 1 + r \cdot \big(\lambda_{r,\theta}(\hat{J} \setminus \hat{u}_1, \hat{\pi} \setminus \hat{u}_1) - \theta \cdot \deg_{\hat{J}}(\hat{u}_1)\big). \tag{6.12}$$

□

It remains to prove Lemma 6.8.

Proof of Lemma 6.8. To simplify the notation we drop the subscripts from $\lambda_{r,\theta}$ and $\mu_{r,\theta}$ and write λ and μ instead. For the reader's convenience, Figure 6.1 illustrates the notations used throughout the proof. Let

$$\begin{aligned}\varepsilon = \varepsilon(F, r, \theta) = \min\Big\{ |\lambda(H_1, \pi_1) - \lambda(H_2, \pi_2)| \;\Big|\; &(H_1, \pi_1), (H_2, \pi_2) \in \mathcal{S}(F) \\ &\wedge \lambda(H_1, \pi_1) \neq \lambda(H_2, \pi_2) \Big\}\end{aligned} \tag{6.13}$$

and

$$v_{\max} = v_{\max}(F, r, \theta) = r^{(v(F)r/\varepsilon + 1)|\mathcal{S}(F)| + 2} \cdot v(F) + r \tag{6.14}$$

We prove the lemma by induction on the number of steps in the game. We show that the statement about graphs $(H, \pi) \in \mathcal{S}(F)$ is true as long as the currently revealed graph G_i does not contain an r-matched subgraph K' with $v(K') \leq v_{\max}$ and $\mu(K') < 0$. Once such a subgraph K' appears we are done as it will remain in the game to the end.

After the first step Painter has assigned a color to r vertices and the inequality (6.10) is trivially satisfied: In each color we only have a single

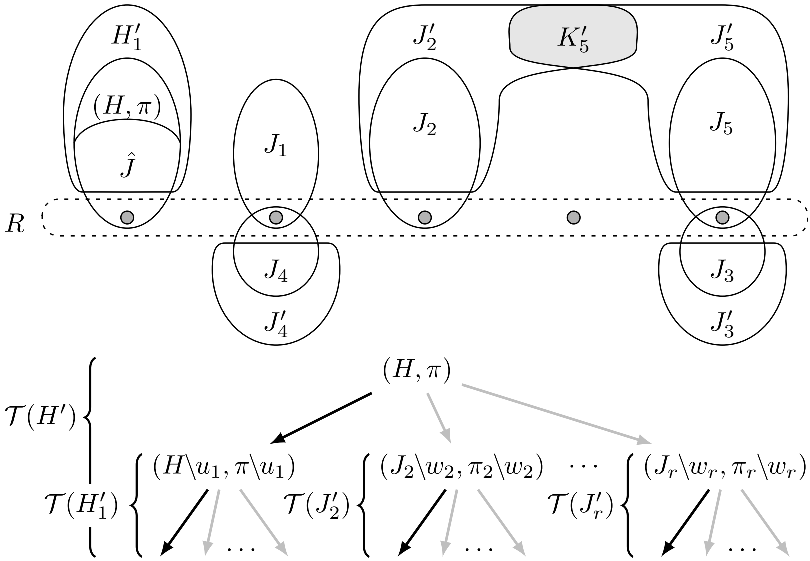


Figure 6.1: Notations used in the proof of Lemma 6.8. The arcs of $\mathcal{T}(H')$ drawn grey are either grey or red in the proof.

vertex, which has a λ-value of 1 according to (6.4). Each of these vertices is contained in the r-matched graph induced by the first r-set, whose μ-value is at most 1, see (6.6).

Consider now an arbitrary step of the game, and denote with $M_1, \dots, M_r$ the matchings Painter considered in this step, where w.l.o.g. M_1 is the matching Painter chose. Assume that M_1 completed a monochromatic copy of $(H \setminus u_1, \pi \setminus u_1)$ to a copy of (H, π) (where u_1 denotes the first vertex of π). Let $\hat{J}$ be some graph in $\arg\min_{J \subseteq H, u_1 \in J} \lambda(J, \pi|_J)$, and note that M_1 also closed a copy of $\hat{J}$. For $1 \leq i \leq r$, let $(J_i, \pi_i) := J(M_i)$ as in the definition of our strategy after (6.9). By definition (J_1, π_1) minimizes $\lambda()$ over all monochromatic ordered graphs in $\mathcal{S}(F)$ that are closed by M_1, see (6.8). Furthermore, since Painter preferred M_1 over the alternatives we have $\lambda(J_1, \pi_1) \geq \lambda(J_i, \pi_i)$, $2 \leq i \leq r$. Taken together it follows that

$$\lambda(\hat{J}, \pi|_{\hat{J}}) \geq \lambda(J_1, \pi_1) \geq \lambda(J_i, \pi_i) \quad \text{for } 2 \leq i \leq r. \tag{6.15}$$

Note that H, $\hat{J}$ or J_1 might be the same graph.

For $1 \leq i \leq r$, let w_i denote the youngest vertex of J_i according to π_i; i.e., $\pi_i = (w_i, \dots)$. Again by the definition of our strategy the graphs (J_i, π_i) minimize $\lambda()$ among all graphs that are closed by M_i, $1 \leq i \leq r$. As for each index i the family of these graphs is subgraph-closed in the sense of Lemma 6.9, it follows that

$$\lambda(J_i, \pi_i) = 1 + r\Big(\lambda(J_i \setminus w_i, \pi_i \setminus w_i) - \theta \cdot \deg_{J_i}(w_i)\Big).$$

Similarly, Lemma 6.9 also yields that

$$\lambda(\hat{J}, \pi|_{\hat{J}}) = 1 + r\Big(\lambda(\hat{J} \setminus u_1, \pi|_{\hat{J} \setminus u_1}) - \theta \cdot \deg_{\hat{J}}(u_1)\Big).$$

Applying these transformations to equation (6.15) yields that for $1 \leq i \leq r$

$$\lambda(\hat{J} \setminus u_1, \pi|_{\hat{J} \setminus u_1}) - \theta \cdot \deg_{\hat{J}}(u_1) \geq \lambda(J_i \setminus w_i, \pi_i \setminus w_i) - \theta \cdot \deg_{J_i}(w_i). \tag{6.16}$$

The copy of $(H \setminus u_1, \pi \setminus u_1)$ on the board is monochromatic and by induction must be contained in some r-matched graph H_1' satisfying equation (6.10), i.e.

$$\mu(H_1') \leq \lambda(H \setminus u_1, \pi \setminus u_1) \ . \tag{6.17}$$

Similarly, the copies of $(J_i \setminus w_i, \pi_i \setminus w_i)$ that are completed to copies of (J_i, π_i), $2 \leq i \leq r$ on the board are also monochromatic, and hence they are contained in r-matched graphs $J_2', \ldots, J_r'$ with

$$\mu(J_i') \leq \lambda(J_i \setminus w_i, \pi_i \setminus w_i) \quad \text{for } 2 \leq i \leq r. \tag{6.18}$$

By induction all these graphs contain at most $v_{\max}$ vertices. We can also assume that $\mu(H_1')$ and $\mu(J_2'), \ldots, \mu(J_r')$ are all non-negative, as otherwise we have found a graph K' with $\mu(K') < 0$ and $v(K') \leq v_{\max}$ and are done. We will argue later that if the μ-values under consideration are indeed non-negative, even stronger bounds on the number of vertices hold; specifically, that

$$\begin{aligned} v(H_1') &< v_{\max}/r - 1 \\ v(J_i') &< v_{\max}/r - 1 \quad \text{for } 2 \leq i \leq r. \end{aligned} \tag{6.19}$$

We now construct an r-matched graph H' for (H, π) satisfying the conditions of the lemma. Denote with E_1 a set of edges that completes the considered copy of $(H \setminus u_1, \pi \setminus u_1)$ to a copy of (H, π) (where the vertex corresponding to u_1 is in R and $|E_1| = \deg_H(u_1)$). Similarly, for $2 \leq i \leq r$ denote with E_i a set of edges that completes the considered copy of $(J_i \setminus w_i, \pi_i \setminus w_i)$ to a copy of (J_i, π_i) (where the vertex corresponding to w_i is in R and $|E_i| = \deg_{J_i}(w_i)$).

Let H' be the r-matched graph obtained by the union of H_1', J_i' and R together with all the edges in E_i, $1 \leq i \leq r$. Formally, we set

$$\begin{aligned} V(H') &:= R \cup V(H_1') \cup \bigcup_{i=2}^{r} V(J_i') \\ E(H') &:= E(H_1') \cup \bigcup_{i=2}^{r} E(J_i') \cup \bigcup_{i=1}^{r} E_i \\ \mathcal{K}(H') &:= \{R\} \cup \mathcal{K}(H_1') \cup \bigcup_{i=2}^{r} \mathcal{K}(J_i') \ . \end{aligned}$$

This is again a well-defined r-matched graph: All r-sets in H_1', J_i' ($2 \leq i \leq r$) and $\{R\}$ are also r-sets of the current game board. As such they are either equal or disjoint. Further $V(H')$ is indeed the union of all r-sets in $\mathcal{K}(H')$, and contains the endpoints of all edges in $E(H')$.

Note that by (6.19) it follows that $v(H') < v_{\max}$.

The r-matched graphs $H_1', J_2', \ldots, J_r'$ are all formed by r-sets that appeared before R in the process and are therefore vertex-disjoint from R. In particular, they do not contain any edges from $E_1, \ldots, E_r$. Furthermore, the sets

$E_1, \ldots, E_r$ are pairwise disjoint: if two such sets E_{i_1}, E_{i_2} involve the same vertex from R, then together with this vertex they complete monochromatic copies of graphs $(J_i \setminus w_i, \pi_i \setminus w_i)$ in two *different* colors to copies of (J_i, π_i); i.e., the endpoints of the edges in E_{i_1}, E_{i_2} outside R are in two different colors and are therefore distinct.

We define the r-matched graphs

$$K'_i = J'_i \cap \Big(H'_1 \cup \bigcup_{j=2}^{i-1} J'_j\Big) \quad \text{for } 2 \leq i \leq r.$$

With the above observations and the definition of $\mu()$ in (6.6) we obtain that

$$\begin{aligned} \mu(H') = 1 &+ \mu(H'_1) - \theta \cdot \deg_H(u_1) \\ &+ \sum_{i=2}^{r} \big(\mu(J'_i) - \theta \cdot \deg_{J_i}(w_i)\big) - \sum_{i=2}^{r} \mu(K'_i). \end{aligned} \tag{6.20}$$

We can assume that all $\mu(K'_i)$ are non-negative, because if this is not the case we have found a graph K' with $\mu(K') < 0$ and $v(K') \leq v_{\max}$ and are done. With this observation and (6.17), (6.18) we obtain that

$$\begin{aligned} \mu(H') \leq 1 &+ \lambda(H \setminus u_1, \pi \setminus u_1) - \theta \cdot \deg_H(u_1) \\ &+ \sum_{i=2}^{r} \big(\lambda(J_i \setminus w_i, \pi_i \setminus w_i) - \theta \cdot \deg_{J_i}(w_i)\big). \end{aligned}$$

Combining this with equation (6.16) yields

$$\begin{aligned} \mu(H') \leq 1 &+ \lambda(H \setminus u_1, \pi \setminus u_1) - \theta \cdot \deg_H(u_1) \\ &+ (r-1) \cdot \Big(\lambda(\hat{J} \setminus u_1, \pi|_{\hat{J}\setminus u_1}) - \theta \cdot \deg_{\hat{J}}(u_1)\Big). \end{aligned} \tag{6.21}$$

By Lemma 6.9 and our choice of $\hat{J}$ the right hand side of the above equation equals $\lambda(H, \pi)$ as defined in (6.4); i.e., we have

$$\mu(H') \leq \lambda(H, \pi)$$

as desired.

It remains to prove that equation (6.19) holds. It suffices to show that given $\mu(H') \geq 0$ we have $v(H') \leq v_{\max}/r - 1$.

In the above argument we constructed H' from copies of H'_1 and J'_i, or in other words from graphs constructed equivalently to H' in prior steps of the induction from $(H \setminus u_1, \pi \setminus u_1)$ and $(J_i \setminus w_i, \pi_i \setminus w_i)$. To analyze this construction we associate it with an edge-colored directed rooted tree $\mathcal{T}(H')$ (cf. Figure 6.1). The vertices of $\mathcal{T}(H')$ correspond to monochromatic copies of graphs from $\mathcal{S}(F)$ on the board of the game (the same copy may appear as a vertex multiple times). If (H, π) consists of a single vertex, then $\mathcal{T}(H')$ consists just of the copy of (H, π) as the root. If this is not the case, then $\mathcal{T}(H')$ consists of the copy of (H, π) as the root vertex joined to r subtrees $\mathcal{T}(H'_1)$ and $\mathcal{T}(J'_i)$, $2 \leq i \leq r$. The subtree $\mathcal{T}(H'_1)$ is connected to the root by a black arc and every $\mathcal{T}(J'_i)$ is connected to the root by either a grey or red arc according to the following criterion: Each such arc corresponds to an instance of the inequalities in (6.15) somewhere along the induction. The arc is grey if both inequalities are tight, i.e., if $\lambda(\hat{J}, \pi|_{\hat{J}}) = \lambda(J_i, \pi_i)$. If on the other hand at least one of the inequalities is strict, i.e., if $\lambda(\hat{J}, \pi|_{\hat{J}}) > \lambda(J_i, \pi_i)$, then the arc is red. All arcs are oriented away from the root. Note that $\mathcal{T}(H')$ captures only the logical structure of the inductive history of H'. Overlappings (captured by the graphs K'_i in (6.20)) are completely ignored.

Every red arc of $\mathcal{T}(H')$ corresponds to a strict inequality in (6.15). In this case, as a consequence of Lemma 6.9, equation (6.16) is also strict, with a difference of at least ε/r (cf. (6.13)) between the right and left side. Consequently, each red arc contributes a term of $-\varepsilon/r$ to the right side of (6.21) in the corresponding induction step. Accumulating these terms along the induction yields that

$$\mu(H') \leq \lambda(H, \pi) - \ell(H') \cdot \varepsilon/r, \tag{6.22}$$

where $\ell(H')$ denotes the number of red arcs in $\mathcal{T}(H')$.

Note that $\lambda(H, \pi) \leq v(F)$ for all $(H, \pi) \in \mathcal{S}(F)$. Thus if $\mu(H') \geq 0$, then by (6.22) the tree $\mathcal{T}(H')$ has at most $\lambda(H, \pi)r/\varepsilon \leq v(F)r/\varepsilon$ many red arcs. We will show that, due to our tie-breaking rule involving the auxiliary graph $\mathcal{G}'$, this bound on the number of red arcs implies the claimed bound of $v_{\max}/r - 1$ on the number of vertices of H'. To that end, we first show that if two vertices of $\mathcal{T}(H')$ are connected by a (directed, i.e. descending) path P that contains no red arcs, then these two vertices correspond to copies of *different* ordered graphs (H_1, π_1), $(H_2, \pi_2) \in \mathcal{S}(F)$.

Consider such a walk between two vertices (H_1, π_1) and (H_2, π_2). We can map P to a directed walk P' in $\mathcal{G}'$ as follows. The initial vertex of P' is (H_1, π_1). For each black arc in P from a copy of some $(H, \pi) \in \mathcal{S}(F)$ to a copy of $(H \setminus u_1, \pi \setminus u_1)$ we extend P' by an arc from (H, π) to $(H \setminus u_1, \pi \setminus u_1)$. This arc exists in $\mathcal{G}'$ by subgraph containment. For each grey arc in P from a copy of some graph (H, π) to a copy of some graph $(J_i \setminus w_i, \pi_i \setminus w_i)$ for some $2 \leq i \leq r$, we have

$$\lambda(H, \pi) \geq \lambda(\hat{J}, \pi|_{\hat{J}}) = \lambda(J_1, \pi_1) = \lambda(J_i, \pi_i). \tag{6.23}$$

In $\mathcal{G}'$ we can then walk between the first two graphs in the above equation (assuming that they are different) because the second is contained in the first. Further we can walk from the second to the third because $(J_1, \pi_1) = J(M_1)$, and therefore by definition it must be lower in the ordering than $(\hat{J}, \pi|_{\hat{J}})$, see the text just after (6.9). The walk between the last two graphs in (6.23) is possible because Painter chose the matching M_1, and by our tie-breaking criterion this means that $(J_1, \pi_1) = J(M_1)$ is higher in the ordering than $(J_i, \pi_i) = J(M_i)$. The last arc between (J_i, π_i) and $(J_i \setminus w_i, \pi_i \setminus w_i)$ is in $\mathcal{G}$ by subgraph containment. We extend P' by all these arcs as well (if any two subsequent graphs in this walk are the same, then the corresponding step in the walk is skipped). Proceeding in this manner we obtain a directed walk P' in $\mathcal{G}'$ from (H_1, π_1) to (H_2, π_2). As $\mathcal{G}'$ is acyclic we must have $(H_1, \pi_1) \neq (H_2, \pi_2)$.

It follows that a (directed) path in $\mathcal{T}(H')$ that contains no red arcs has at most $|\mathcal{S}(F)|$ many vertices. Since in total we have at most $v(F)r/\varepsilon$ many red arcs in $\mathcal{T}(H')$, it follows that the depth of $\mathcal{T}(H')$ is bounded by

$$(v(F)r/\varepsilon + 1)|\mathcal{S}(F)|\ ,$$

and that consequently

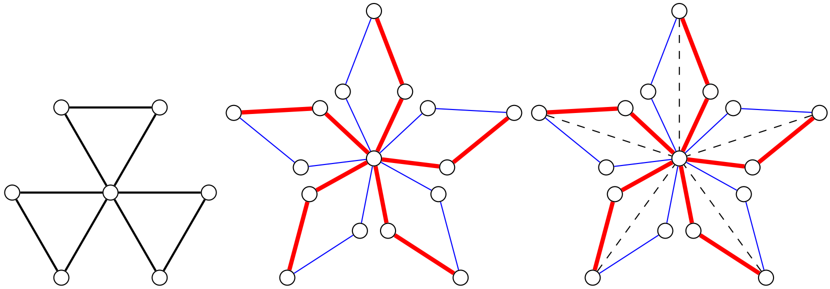
$$v\big(\mathcal{T}(H')\big) \leq 1 + r + r^2 + \cdots + r^{(v(F)r/\varepsilon+1)|\mathcal{S}(F)|} \leq r^{(v(F)r/\varepsilon+1)|\mathcal{S}(F)|+1}.$$

Since each vertex of $\mathcal{T}(H')$ corresponds to at most $v(F)$ vertices of H' we finally obtain that

$$v(H') \leq r^{(v(F)r/\varepsilon+1)|\mathcal{S}(F)|+1} \cdot v(F) \stackrel{(6.14)}{=} v_{\max}/r - 1.$$

$\square$

Remark 6.10. *The reader might wonder where exactly an attempt to extend the* edge *Achlioptas lower bound proof in the same way fails. The issue arises with the definition of the graphs K_i' that capture possible overlaps of the r-matched graphs $H_1', J_2', \ldots, J_r'$. In the edge case it is not possible to define these in such a way that the analogue of* (6.20) *holds.*



(a) The graph $C_{3,3}$ (b) The graph $C^{2+}_{3,3}$ (c) The graph $C^{2*}_{3,3}$, with its center star dashed

Figure 6.2: A graph with different thresholds for the Achlioptas and the balanced Ramsey game.

6.5 The edge case

In this section we prove Theorem 6.1, our separation result for the edge case.

As already mentioned, it is not hard to see that the Achlioptas game and the balanced Ramsey game have different thresholds for certain forests. The simplest example is the case where F is the star with three rays and $r = 2$: By the pigeon-hole principle, in the balanced Ramsey game the player will lose the game as soon as the board contains a star with five rays, which by a standard result a.a.s. happens after $\Theta(n^{2-6/5}) = \Theta(n^{4/5})$ many steps (see e.g. [JŁR00, Section 3.1]). Thus the threshold of the balanced Ramsey game is bounded from above by $n^{4/5}$. In the Achlioptas game with the same parameters on the other hand, stars on 5 edges are not an issue, as typically the player can simply choose not to pick more than 2 edges out of each such star. Specifically, the results of [MST11] yield a strictly higher threshold of $n^{6/7}$ for the Achlioptas game.

As it turns out, similar pigeon-hole problems as in the star example may arise for more complex graphs as well. The simplest such example is given by the graph $C_{3,3}$ consisting of 3 triangles joined at one vertex, see Figure 6.2(a). The results of [MST11] yield a threshold of $n^{2-22/35} = n^{1.371\ldots}$ for the Achlioptas game with this graph and $r = 2$. As we will see, the threshold of the corresponding balanced Ramsey game is at most $n^{1.36}$.

The reason is that, regardless of the strategy Painter uses, many copies of the graph $C_{3,3}^{2+}$ colored exactly as in Figure 6.2(b) will appear relatively early in the game. Once all 5 edges drawn dashed in Figure 6.2(c) have appeared in such a copy, by the pigeon-hole principle Painter will have created a monochromatic copy of $F = C_{3,3}$. As $C_{3,3}^{2*}$ has 16 vertices and 25 edges, the upper bound resulting from this argument is $n^{2-16/25} = n^{1.36}$.

We will show that this argument generalizes to any graph F formed by some number of cycles of the same length joined at a common vertex, and to any number $r \geq 2$ of colors.

Definition 6.11. *Let $C_{\ell,k}$ denote the graph obtained by joining k cycles of length ℓ at one common vertex.*

We will prove:

Theorem 6.12. *For all integers $\ell \geq 3$, $k \geq 3$, and $r \geq 2$, the threshold of the the balanced Ramsey game with parameters $C_{\ell,k}$ and r is strictly lower than the threshold of the Achlioptas game with the same parameters.*

We first give an upper bound on the threshold of the balanced Ramsey game with parameters $C_{\ell,k}$ and r. To do so we will use an offline result that is very similar and can be proved completely analogously to [KSS10, Theorem 15] for the Achlioptas case. For any graph F with at least one edge, we let

$$m_2(F) := \max_{H \subseteq F : v(H) \geq 3} \frac{e(H) - 1}{v(H) - 2} \tag{6.24}$$

if $v(F) \geq 3$, and $m_2(F) = 1/2$ otherwise (i.e., if $F = K_2$). By $G^r(n,m)$ we denote a random r-edge-matched graph obtained by sampling a random graph $G(n,m)$ on n vertices with m edges uniformly at random, and then partitioning the m edges into sets of size r uniformly at random (we assume that m is divisible by r). Note that by symmetry the board of the edge Achlioptas or balanced Ramsey game after m/r steps is distributed exactly like $G^r(n,m)$. A *balanced coloring* of $G^r(n,m)$ is an edge-coloring that uses each of the r available colors for exactly one edge in each r-set. Note that in the balanced Ramsey game, the goal is to find such a balanced coloring in an online setting. The following theorem concerns the same problem in an *offline* setting.

Theorem 6.13. *Let F be a fixed graph with at least one edge, and let $c : E(F) \to \{1, \ldots, r\}$ be an arbitrary edge-coloring of F. There exist positive constants $C = C(F,r)$ and $a = a(F,r)$ such that for $m \geq Cn^{2-1/m_2(F)}$*

with $m \ll n^2$, a.a.s. every balanced coloring of $G^r(n,m)$ contains at least $an^{v(F)}(m/n^2)^{e(F)}$ many copies of F colored as specified by c. $\square$

We now prove the desired upper bound on the balanced Ramsey threshold for the graphs $C_{\ell,k}$.

Lemma 6.14. *For all integers $\ell \geq 3$, $k \geq 3$, and $r \geq 2$, the threshold for the balanced Ramsey game with parameters $C_{\ell,k}$ and r is at most*

$$N_{UB\text{-}bal}(\ell,k,r,n) := n^{2-\frac{(r(\ell-2)+1)(r(k-1)+1)+1}{(r(\ell-1)+1)(r(k-1)+1)}}.$$

Proof. Consider the graph obtained by joining one endpoint of r paths of length $\ell-1$ in one common vertex and the other endpoint of each in a second common vertex. We call this graph a *petal* and the two vertices in which all paths meet the *endpoints* of the petal. We will refer to the non-edge connecting the two endpoints of a petal as the *missing edge* of that petal. Let $C_{\ell,k}^{r+}$ denote the graph obtained by joining one endpoint of $k^* := r(k-1)+1$ many petals at a common vertex. We say that a copy of $C_{\ell,k}^{r+}$ on the game board is *properly colored* if the two endpoints of each of its petals are connected by a path (of length $\ell-1$) in each color. See Figure 6.2(b) for an example of a properly colored $C_{3,3}^{2+}$. The *center star* of a copy of $C_{\ell,k}^{r+}$ is the graph obtained as the union of all missing edges of the petals of $C_{\ell,k}^{r+}$. We denote with $C_{\ell,k}^{r*}$ the union of $C_{\ell,k}^{r+}$ and its center star, cfr. Figure 6.2(c). Clearly, we have

$$\begin{aligned} e(C_{\ell,k}^{r*}) &= e(C_{\ell,k}^{r+}) + k^* = (r(\ell-1)+1)(r(k-1)+1), \\ v(C_{\ell,k}^{r*}) &= v(C_{\ell,k}^{r+}) = (r(\ell-2)+1)(r(k-1)+1)+1. \end{aligned}$$

Let $d^* := e(C_{\ell,k}^{r*})/v(C_{\ell,k}^{r*})$, and note that $N_{\text{UB-bal}}(\ell,k,r,n) = n^{2-1/d^*}$.

It is not hard to check that $m_2(C_{\ell,k}^{r+}) = (r(\ell-1)-1)/(r(\ell-2))$ (the maximum in (6.24) is attained by a single petal of $C_{\ell,k}^{r+}$), and it is also quite straightforward to verify that this quantity is strictly less than d^*.

Let now $N \gg N_{\text{UB-bal}} = n^{2-1/d^*}$ with $N \ll n^2$ be given, and assume w.l.o.g. that N is even. Set $p := nrN/n^2$. Observing that $N \gg n^{2-1/d^*} \geq n^{2-1/m_2(C_{\ell,k}^{r+})}$, we obtain with Theorem 6.13 that a.a.s., after $N/2$ steps of the balanced Ramsey game the board contains

$$a(C_{\ell,k}^{r+}, r) \cdot n^{v(C_{\ell,k}^{r+})} p^{e(C_{\ell,k}^{r+})} 2^{-e(C_{\ell,k}^{r+})} =: M'$$

many properly colored copies of $C_{\ell,k}^{r+}$, regardless of Painter's strategy. Furthermore, the expected number of copies of $C_{\ell,k}^{r+}$ (ignoring any coloring) in which at least one edge of the center star is already present after $N/2$ steps is $\mathcal{O}(M'p) = o(M')$. It follows with Markov's inequality that after $N/2$ steps a.a.s. there are at least $M := 0.99M'$ properly colored copies of $C_{\ell,k}^{r+}$ such that in each of these, none of the edges of the center star is already present.

Let $C_{\ell,k}^{r+} \cup_J C_{\ell,k}^{r+}$ denote the union of two copies of $C_{\ell,k}^{r+}$ which intersect in a graph J and whose (missing) center stars intersect in a nonempty graph J_S. Further let $J^* := J \cup J_S$. Let the random variable M_J denote the number of copies of $C_{\ell,k}^{r+} \cup_J C_{\ell,k}^{r+}$ (ignoring any coloring) contained in the game board after the first $N/2$ steps. We have

$$\begin{aligned}\mathbb{E}[M_J] &= \Theta(n^{2v(C_{\ell,k}^{r+})-v(J)}p^{2e(C_{\ell,k}^{r+})-e(J)}) \\ &= \Theta(n^{2v(C_{\ell,k}^{r+})}p^{2e(C_{\ell,k}^{r+})})n^{-v(J)}p^{-e(J)} \\ &= \Theta(M^2)n^{-v(J^*)}p^{-e(J^*)+e(J_S)},\end{aligned} \tag{6.25}$$

where in the last step we used that $e(J^*) = e(J) + e(J_S)$.

Note that $C_{\ell,k}^{r*}$ is a balanced graph, i.e. for all subgraphs $H \subseteq C_{\ell,k}^{r*}$ with $v(H) \geq 1$ we have $e(H)/v(H) \leq e(C_{\ell,k}^{r*})/v(C_{\ell,k}^{r*}) = d^*$. This holds in particular also for $H = J^*$. As $p \gg n^{-1/d^*}$, it follows that $n^{v(J^*)}p^{e(J^*)} = \omega(1)$. Hence by Markov's inequality we obtain from (6.25) that a.a.s.

$$M_J = o(M^2)p^{e(J_S)} \tag{6.26}$$

(i.e., for an appropriate function $f(n) = o(1)$ a.a.s. it holds that $M_J \leq f(n)M^2p^{e(J_S)}$).

For the remaining $N/2$ steps of the game we condition on having at least M properly colored copies of $C_{\ell,k}^{r+}$ whose center star edges are not already present, and on M_J being as above for all $J \subseteq C_{\ell,k}^{r+}$. As the number of graphs J is a constant depending only on k, ℓ, and r, a.a.s all these properties hold simultaneously after $N/2$ steps. Using the second moment method, we will show that in the remaining $N/2$ steps, a.a.s. in at least one of the properly colored copies of $C_{\ell,k}^{r+}$ all edges of the center star will appear. Clearly, this then forces Painter to complete a monochromatic copy of $C_{\ell,k}$ by the pigeon-hole principle.

Fix a family of exactly M properly colored copies of $C_{\ell,k}^{r+}$ (say the lexicographally first ones; w.l.o.g. M is an integer), and let $S_1, \ldots, S_M$ denote the (not necessarily distinct) center stars of these copies. For each S_i let Z_i denote the indicator random variable for the event that the k^* edges of S_i will appear in the remaining $N/2$ steps of the game. Let Z denote the sum over all Z_i. As the $rN/2$ random edges revealed in the second half of the game are distributed uniformly among the $\binom{n}{2} - rN/2$ edges never seen before, we have

$$\mathbb{E}[Z_i] = \frac{\binom{\binom{n}{2}-rN/2-k^*}{rN/2-k^*}}{\binom{\binom{n}{2}-rN/2}{rN/2}} = \Theta(p^{k^*})$$

for all i, and hence

$$\mathbb{E}[Z] = \Theta(Mp^{k^*}) = \Theta(n^{v(C^{r+}_{\ell,k})} p^{e(C^{r+}_{\ell,k})+k^*}) = \Theta\big(n^{v(C^{r*}_{\ell,k})} p^{e(C^{r*}_{\ell,k})}\big).$$

By our choice of N this quantity is $\omega(1)$.

It remains to establish concentration of Z via the second moment method — it then follows that $Z \geq 1$ a.a.s., which as discussed implies that Painter loses the game. We have

$$\begin{aligned}\mathrm{Var}[Z] &= \sum_{i,j=1}^{M} \big(\mathbb{E}[Z_i Z_j] - \mathbb{E}[Z_i]\mathbb{E}[Z_j]\big) \leq \sum_{\substack{J \subseteq C^{r+}_{\ell,k} \\ e(J_S) \geq 1}} M_J \cdot \Theta(p^{2k^*-e(J_S)}) \\ &\overset{(6.26)}{=} \sum_{\substack{J \subseteq C^{r+}_{\ell,k} \\ e(J_S) \geq 1}} o(M^2) p^{2k^*} = o(\mathbb{E}[Z]^2).\end{aligned}$$

The last equality follows from the fact that the number of possible choices for J is a constant depending only on k, ℓ and r. This concludes the proof. □

We conclude the proof of Theorem 6.12 by deriving a lower bound on the Achlioptas threshold for the graphs $C_{\ell,k}$ from the general formula given in [MST11].

Lemma 6.15. *For all integers $\ell \geq 3$, $k \geq 3$, and $r \geq 2$, the threshold for the Achlioptas game with parameters $C_{\ell,k}$ and r is at least*

$$N_{LB\text{-}Achl}(\ell, k, r, n) = n^{2 - \frac{(r(l-2)+1)(r^k-1)+r-1}{(r(l-1)+1)(r^k-1)}}.$$

Proof. For the reader's convenience we reproduce the general edge Achlioptas threshold formula here. For notational details we refer to [MST11].

For any nonempty edge-ordered graph (H_1, π), $\pi = (e_1, \ldots, e_h)$, any sequence of subgraphs $H_2, \ldots, H_h \subseteq H_1$ with $H_i \subseteq H_1 \setminus \{e_1, \ldots, e_{i-1}\}$ and $e_i \in H_i$ for all $2 \leq i \leq h$, and any integer $r \geq 2$, define coefficients $c_i = c_i((H_1, \pi), H_2, \ldots, H_h, r)$ recursively by

$$
\begin{aligned}
c_1 &:= r, \\
c_i &:= (r-1) \cdot \sum_{j=1}^{i-1} c_j \mathbf{1}_{\{e_i \in H_j\}}, \quad 2 \leq i \leq h,
\end{aligned}
$$

(where $\mathbf{1}_{\{e_i \in H_j\}} = 1$ if $e_i \in H_j$ and $\mathbf{1}_{\{e_i \in H_j\}} = 0$ otherwise), and set

$$
d^{r*}(H_1, \pi) := \max_{\substack{H_2, \ldots, H_h \\ \forall i \geq 2:\ H_i \subseteq H_1 \setminus \{e_1, \ldots, e_{i-1}\} \ \wedge\ e_i \in H_i}} \frac{1 + \sum_{i=1}^{h} c_i (e(H_i) - 1)}{2 + \sum_{i=1}^{h} c_i (v(H_i) - 2)}. \tag{6.27}
$$

Furthermore, we set for any integer $r \geq 2$ and any nonempty graph F

$$
m^{r*}(F) := \min_{\pi \in \Pi(E(F))} \max_{H_1 \subseteq F} d^{r*}(H_1, \pi|_{H_1}).
$$

The threshold of the Achlioptas game with parameters F and r is then given by $N_0(F, r, n) = n^{2 - 1/m^{r*}(F)}$.

We now prove that $m^{r*}(C_{\ell,k})$ is bounded from below as claimed in the lemma. Let $\pi = (e_1, \ldots, e_h)$ be an arbitrary permutation of the edges of $C_{\ell,k}$. Denote with $e_{t_1}, \ldots, e_{t_k}$ the first edge in each of the k cycles of $C_{\ell,k}$ according to π, in order of their appearance in π. (Thus in particular $e_{t_1} = e_1$.) Let $C_1, \ldots, C_k$ the corresponding cycles in $C_{\ell,k}$, i.e. $e_{t_i} \in C_i$ for all i. Choose

$$
H_i = \begin{cases} e_i & i \notin \{t_1, \ldots, t_k\} \\ \bigcup_{j \geq i}^{k} C_j & i = t_j \end{cases}.
$$

Note that this choice is compatible with the requirements of (6.27). This yields

$$
\begin{aligned}
e(H_{t_j}) &= (k - j + 1)\ell \\
v(H_{t_j}) &= (k - j + 1)(\ell - 1) + 1,
\end{aligned}
$$

and

$$
c_{t_j} = \begin{cases} r & j = 1 \\ (r-1) r^{j-1} & j \neq 1. \end{cases}
$$

Note that the coefficients c_i for $i \notin \{t_1, \ldots, t_k\}$ are not required, as both $e(H_i) - 1$ and $v(H_i) - 2$ are 0. It is somewhat tedious but straightforward to verify that

$$\frac{1 + \sum_{i=1}^{h} c_i \big(e(H_i) - 1\big)}{2 + \sum_{i=1}^{h} c_i \big(v(H_i) - 2\big)} = \frac{(r(l-1)+1)(r^k - 1)}{(r(l-2)+1)(r^k - 1) + r - 1}.$$

As this holds for any edge ordering $\pi \in \Pi(E(C_{\ell,k}))$, we readily obtain the desired lower bound

$$\begin{aligned} m^{r*}(C_{\ell,k}) &= \min_{\pi \in \Pi(E(C_{\ell,k}))} \max_{H_1 \subseteq C_{\ell,k}} d^{r*}(H_1, \pi|_{H_1}) \\ &\geq \frac{(r(l-1)+1)(r^k - 1)}{(r(l-2)+1)(r^k - 1) + r - 1}. \end{aligned} \qquad \square$$

Theorem 6.12 now follows, after some calculation, from Lemmas 6.14 and 6.15.

6.6 Proofs for the vertex Achlioptas upper bound

6.6.1 Upper bound for the Achlioptas game

In this section we prove Theorem 6.3. The proof here is an adaptation to the vertex case of the corresponding edge-case proof in [MST11]. We make use of Theorem 6.4 (ignoring the last sentence in its statement; this is restated and proved as Lemma 6.21 below) and a technical lemma (Lemma 6.19 below), both of which are proved in the next section.

Before we start we wish to present the following adaptation to r-matched graphs of Bollobás' classical small subgraphs result [Bol81].

Theorem 6.16. *Let $r \geq 2$ be a fixed integer, and let F be a fixed r-matched graph with at least one edge. Define*

$$m^r(F) := \max_{\substack{H \subseteq F:\\ \kappa(H) > 0}} \frac{e(H)}{\kappa(H)}.$$

Then the threshold for the appearance of F in $G^r(n,p)$ is

$$p_0(n) = n^{-1/m^r(F)}.$$

Further, if $p \gg n^{-1/m^r(F)}$ we have that the number of copies of F in $G^r(n,p)$ is a.a.s.

$$\Theta(n^{\kappa(F)}p^{e(F)}).$$

One can prove this by an easy application of the first and second moment method. We do not require this result, but it is useful to gain a better intuition for our proof. Note that we could state the first part of Theorem 6.16 equivalently as follows.

Theorem 6.17. *Let $r \geq 2$ be a fixed integer, and let F be a fixed r-matched graph with at least one edge. Let $\theta' = \theta'(F,r)$ be the unique solution of*

$$\min_{H \subseteq F} \mu_{r,\theta}(H) \stackrel{!}{=} 0,$$

where $\mu_{r,\theta}$ is defined in (6.6). Then the threshold for the appearance of F in $G^r(n,p)$ is

$$p_0(n) = n^{-\theta'}.$$

Concerning the second part of Theorem 6.16, recall also that for $p = n^{-\theta}$ we have $n^{\kappa(F)}p^{e(F)} = n^{\mu_{r,\theta}(F)}$. The two "dual" formulations of our threshold result in Theorem 6.3 and (the last sentence of) Theorem 6.4 are related to each other similarly as the two statements above.

In order to prove Theorem 6.3 it is not sufficient to consider only a graph F, we additionally need to consider the order in which its vertices are presented to the player. In our proof this is encoded by an ordered graph (F,π). Recall that when we use terms such as first or last for the vertices of F we mean this with respect to the order in which they are presented to the player. In that context, for $\pi = (u_1, \ldots, u_f)$ the last vertex is u_1 and the first u_f.

If as an adversary we wanted to force the player to create a copy of F, we would wish to be able to present r copies of F_- (F without the last vertex), an additional r-set, and edges such that choosing any of the vertices in the r-set completes a copy of F_- to a copy of F. In such a situation the player has no choice and loses. Of course the player could try to avoid creating copies of F_- in the first place, so that this situation does not arise. However, applying the same argument recursively, we could force the creation of r copies of F_- by r^2 copies of F_{2-} (F missing the last 2 vertices), r many r-sets, and all edges necessary to join each of the r^2 vertices in the r-sets to a different copy of F_{2-}, in such a way as to form r^2 many "threats" for the player. The player must choose one vertex in all of these r many r-sets and is thus forced to create r copies of F_-. We can continue this reasoning recursively until at the tail end we have $r^{v(F)-1}$ disjoint r-sets, out of each of which the player is forced to choose 1 vertex. Assuming that such a recursive "history graph" for F appears in the game in the correct order, it would guarantee that the player is left no choice but to create a copy of F. We denote such a construct with F_r^π and formalize its definition below.

In the following, a *grey-black r-matched graph* is a tuple $H = (V, E, \mathcal{K}, B)$, where $(V, E, \mathcal{K})$ is an r-matched graph, and B is a set of vertices containing exactly one vertex from every r-set in $\mathcal{K}$. We interpret B as the set of vertices chosen by the player during the game, and call them the black vertices. The remaining $|\mathcal{K}|(r-1)$ vertices are considered grey (to indicate "presented but not chosen"). Sometimes we ignore the coloring and tacitly identify H with the underlying r-matched graph $(V, E, \mathcal{K})$.

Recall that the board of the game is distributed as a random r-matched graph $G^r(n,p)$, and that we defined its state after $1 \le i \le n/r$ rounds with G_i. For the purpose of this section we additionally require information about which vertices were chosen by the player. To this end, for each $1 \le i \le n/r$, in this section we append to G_i the set B_i of vertices chosen by the player up to round i and consider G_i to be a grey-black r-matched graph.

Definition 6.18. *Let (F, π) be an ordered graph with $\pi = (u_1, \dots, u_f)$. Then we define the* grey-black *r-matched graph F_r^π and a distinguished black copy of (F, π), the* central copy of (F, π) in F_r^π, *recursively as follows*

- *If $v(F) = 1$, then F_r^π consists of one r-set with a distinguished black vertex. This vertex is the central copy of F in F_r^π.*
- *If $v(F) \ne 1$, then F_r^π consists of the disjoint union of r copies of $(F \setminus u_1)_r^{\pi \setminus u_1}$, denoted $F_{-,1}^\pi, \dots, F_{-,r}^\pi$, an additional r-set $(v_1, \dots, v_r)$, and $r \deg_F(u_1)$ many additional edges which for all $1 \le i \le r$ connect v_i to $F_{-,i}^\pi$ and extend the central copy of $(F \setminus u_1, \pi \setminus u_1)$ in $F_{-,i}^\pi$ to a copy of (F, π). The vertex v_1 is chosen as black and the copy of (F, π) containing it is the central copy of (F, π) in F_r^π.*

We refer to the additional r-set in the recursive step as the central r-set *of F_r^π.*

As explained above, if the r-sets of a copy of F_r^π are presented to the player in an ordering such that all r-sets deeper in the recursion are presented before those at lower recursion depths, then the player is forced to create a copy of F. As it turns out, the threshold for this to happen in the game coincides with the threshold for the appearance of F_r^π in the random r-matched graph $G^r(n,p)$ (as an r-matched graph without any ordering or coloring). This last threshold is given by Theorem 6.16.

Note that the ordering π on the vertices of F is crucial. For different choices of π the corresponding grey-black r-matched graphs F_r^π may have very different thresholds for their appearance in $G^r(n,p)$. As the player has no influence over the order in which r-sets are presented to her, the threshold for the game is bounded from above by, and indeed coincides with, the *lowest* threshold for the appearance of F_r^π over all choices of π. I.e., the threshold stated in Theorems 6.3 and 6.4 can alternatively be written as

$$p_0(F,r,n) = \min_{\pi\in\Pi(V(F))} n^{-1/m^r(F_r^\pi)},$$

where $m^r(F_r^\pi)$ is as defined in Theorem 6.16.

At a high level, the proof is an induction over $v(F)$ and mirrors the recursive definition of F_r^π. At each step of the induction we divide the r-sets presented to the player in 2 halves. We let the player play on the first half and by induction we know that a.a.s. she must have created many copies of $(F\setminus u_1)^{\pi\setminus u_1}$ (in the notation of Definition 6.18). Then we let the player play on the second half of the r-sets and argue via first and second moment method that, conditional on a "good" first round, a.a.s. enough r-sets presented in the second round are connected to r copies of $(F\setminus u_1)^{\pi\setminus u_1}$ as in Definition 6.18.

To apply the second moment method we need the following lemma, which essentially states that for $p \gg n^{-\theta'}$, where θ' is defined below, the expected number of copies in $G^r(n,p)$ of any subgraph of F_r^π is $\omega(1)$, cf. the remark after Theorem 6.17.

Lemma 6.19. *Let $r\geq 2$ be an integer, and let (F,π) be a nonempty ordered graph. Let F_r^π be as in Definition* 6.18*, and let $\theta'=\theta'(F,\pi,r)$ be the unique solution of*

$$\min_{H\subseteq F} \lambda_{r,\theta}(H,\pi|_H) \stackrel{!}{=} 0, \tag{6.28}$$

where $\lambda_{r,\theta}()$ is defined in (6.4)*. Then every r-matched subgraph $J\subseteq F_r^\pi$ satisfies*

$$\mu_{r,\theta'}(J)\geq 0,$$

where $\mu_{r,\theta'}()$ is defined in (6.6)*.*

The proof of this lemma is long and technical, and therefore postponed to the next subsection.

The next lemma implements the inductive proof strategy outlined above. The parameter t ensures that we can require inductively that $r \cdot t$ copies of e.g. $(F \setminus u_1)_r^{\pi \setminus u_1}$ evolve into t copies of F_r^π.

Lemma 6.20. *Let $r \geq 2$ be an integer, and let (F, π) be a nonempty ordered graph. Let $t \geq 1$ be an integer, and let $\mathcal{F}_r^\pi := t \cdot F_r^\pi$ denote the disjoint union of t copies of F_r^π. If $1 \gg p \gg n^{-\theta'}$, where $\theta' = \theta'(F, \pi, r)$ is the unique solution of*

$$\min_{H \subseteq F} \lambda_{r,\theta}(H, \pi|_H) \stackrel{!}{=} 0\,,$$

and $\lambda_{r,\theta}()$ is defined in (6.4), then a.a.s. the number of copies of $\mathcal{F}_r^\pi$ (as a grey-black r-matched graph) in $G_{n/r}$ is

$$\Omega\big(n^{\kappa(\mathcal{F}_r^\pi)} p^{e(\mathcal{F}_r^\pi)}\big) \tag{6.29}$$

regardless of the strategy of the player.

Before we prove this lemma, we show how it implies Theorem 6.3.

Proof of Theorem 6.3. By the equivalence stated in Theorem 6.4 (and proved in Lemma 6.21 below), it suffices to prove that for $p \gg n^{-\theta^*(F,r)}$, the player will a.a.s. create a copy of F no matter how she plays. Let $\pi \in \Pi(V(F))$ be an ordering maximizing the right hand side of (6.5) for $\theta = \theta^*(F, r)$, such that $\theta^*(F, r) = \theta'(F, \pi, r)$ for θ' as in Lemmas 6.19 and 6.20. Applying Lemma 6.20 for $t = 1$, we obtain that $G_{n/r}$ a.a.s. contains

$$\Omega(n^{\kappa(F_r^\pi)} p^{e(F_r^\pi)}) \gg n^{\mu_{r,\theta^*}(F)} \geq 1,$$

many copies of F_r^π, where the last inequality follows from Lemma 6.19. The central copy of (F, π) in each of these copies of F_r^π is black, i.e. all its vertices were selected by the player. $\square$

We now prove Lemma 6.20, giving the main inductive argument of our upper bound proof.

Proof of Lemma 6.20. We prove this lemma by induction on $v(F)$ using the second moment method.

To simplify the notation we drop all subscripts r from F_r^π and $\mathcal{F}_r^\pi$.

As a base case for the induction we consider the case of an empty F. The lemma does not apply to this case directly (as θ' is not well-defined), but a statement equivalent to (6.29) still holds and is all that we require for the induction. If F contains no edges then $\mathcal{F}^\pi$ contains no edges either and consists only of $\kappa(\mathcal{F}^\pi)$ many disjoint r-sets. It trivially holds that $G_{n/r}$ contains $\Theta(n^{\kappa(\mathcal{F}^\pi)})$ many copies of $\mathcal{F}^\pi$, regardless of the choice of p.

To discuss the induction step we first introduce some notation. Let $\pi = (u_1, \ldots, u_f)$, $\pi_- = \pi \setminus u_1$ and $F_- = F \setminus u_1$. Further let F_-^π denote the grey-black r-matched graph $(F_-)^{\pi_-}$, and denote by $\mathcal{F}_-^\pi$ the disjoint union of rt copies of F_-^π.

We use a two-round approach for the induction step. In the first round we let the player make her choices for all r-sets in $G_{n/(2r)}$. By the induction hypothesis we obtain a lower bound on the number of copies of $\mathcal{F}_-^\pi$ that the player must have created which holds with high probability. Conditioning on the fact that the bound from the first round holds, we then derive a bound for the number of copies of $\mathcal{F}^\pi$ which the player is forced to create when she is presented the remaining r-sets in $G_{n/r}$.

Note that if F_- is nonempty, then $\theta'(F_-, \pi_-, r) \geq \theta'(F, \pi, r)$ (cf. (6.28)), and we can apply the induction hypothesis for $t \leftarrow r \cdot t$ and $(F, \pi) \leftarrow (F_-, \pi_-)$ to $G_{n/(2r)}$. If F_- is empty we apply the base case of the induction described above.

We have that a.a.s. at least

$$N := cn^{\kappa(\mathcal{F}_-^\pi)} p^{e(\mathcal{F}_-^\pi)} \tag{6.30}$$

copies of $\mathcal{F}_-^\pi$ are created in the first round for some appropriate constant $c > 0$. For the second round we condition on this event (and also on (6.36) below, which is however irrelevant for the time being). We fix a set of exactly N copies of $\mathcal{F}_-^\pi$ (say the N lexicographically first ones), and only consider these throughout the following.

Recall that by the construction given in Definition 6.18 we can extend r copies of F_-^π to one copy of F^π. To do so we add one new r-set and $r \deg_F(u_1)$ edges (u_1 is the last vertex of (F, π)). Each of the r central copies of (F_-, π_-) is connected to a different vertex of the r-set by $\deg_F(u_1)$ edges and becomes a copy of (F, π).

By repeating the above t times in parallel, any t disjoint r-sets presented in the second round together with one of the N copies of $\mathcal{F}_-^\pi$ can be extended to a copy of $\mathcal{F}^\pi$, provided that the required edges appear in the second round of the game.

Let M be the number of possible pairs of t disjoint r-sets and one copy of $\mathcal{F}^\pi_-$. We index these pairs with $i = 1, \dots, M$. For each such pair there may be several possible edge sets which extend the pair to a copy of $\mathcal{F}^\pi$ as described. We fix one arbitrarily and denote this edge set by T_i. We denote with $\mathcal{F}^\pi_i$ the copy of $\mathcal{F}^\pi$ that is created if all edges of T_i appear during the second round. Note that $|T_i| = t \cdot r \deg_F(u_1)$ for all i. By $\mathcal{K}_i$ we denote the family of t disjoint r-sets that belongs to pair i. Note that each such family belongs to N pairs in total.

There are $\Theta(n^t)$ possible ways to choose the r-sets, so we have

$$M = \Theta\big(n^t\big) \cdot N \overset{(6.30)}{=} \Theta\big(n^{t+\kappa(\mathcal{F}^\pi_-)} p^{e(\mathcal{F}^\pi_-)}\big) = \Theta\big(n^{\kappa(\mathcal{F}^\pi)} p^{e(\mathcal{F}^\pi_-)}\big). \tag{6.31}$$

For $i = 1, \dots, M$ we define the indicator variable Z_i for the event that T_i is contained in $G_{n/r}$. Set $Z = \sum_{i=1}^M Z_i$. Note that Z is a lower bound on the number of copies of $\mathcal{F}^\pi$ created during the second round.

For each Z_i we have

$$\mathbb{E}[Z_i] = \mathbb{P}[Z_i = 1] = p^{rt \deg_F(u_1)}. \tag{6.32}$$

For the expected value of Z, conditioned on (6.30), we thus obtain

$$\mathbb{E}[Z] = \sum_{i=1}^{M} \mathbb{E}[Z_i] \overset{(6.32)}{=} M p^{rt \deg_F(u_1)} \overset{(6.31)}{=} \Theta\big(n^{\kappa(\mathcal{F}^\pi)} p^{e(\mathcal{F}^\pi)}\big). \tag{6.33}$$

To apply the second moment method, we need to bound the variance of Z. Denote by $I \subseteq \{1, \dots, M\}^2$ the set of pairs of indices (i, j) such that $T_i \cap T_j \neq \emptyset$. For $(i, j) \in I$ let $\kappa_{ij} = \mathcal{K}_i \cap \mathcal{K}_j$ and $t_{ij} = |T_i \cap T_j|$. For such pairs of indices we have

$$\mathbb{E}[Z_i Z_j] = p^{2rt \deg_F(u_1) - t_{ij}}.$$

For indices i, j with $T_i \cap T_j = \emptyset$ on the other hand Z_i and Z_j are independent and can be dropped from the variance calculation. We obtain

$$\begin{aligned} \mathrm{Var}[Z] &= \sum_{i,j=1}^{M} \Big(\mathbb{E}[Z_i Z_j] - \mathbb{E}[Z_i]\mathbb{E}[Z_j]\Big) \leq \sum_{(i,j) \in I} \mathbb{E}[Z_i Z_j] \\ &= \sum_{(i,j) \in I} p^{2rt \deg_F(u_1) - t_{ij}}. \end{aligned} \tag{6.34}$$

Let $\mathcal{J} \subseteq \mathcal{F}^\pi$ be a subgraph that contains at least one of the t central r-sets of the t copies of F^π in $\mathcal{F}^\pi$. Denote with $\mathcal{K}_\mathcal{J}$ the family of these central r-sets in $\mathcal{J}$, and with $\kappa_\mathcal{J} := |\mathcal{K}_\mathcal{J}|$ their number. Let $\mathcal{J}_-$ be the r-matched graph obtained from $\mathcal{J}$ by removing $\mathcal{K}_\mathcal{J}$ and all incident edges. Let $T_\mathcal{J}$ be the graph induced by the edges of $\mathcal{J}$ between $\mathcal{K}_\mathcal{J}$ and $\mathcal{J}_-$.

Let $M_\mathcal{J}$ denote the number of pairs (i, j) for which the intersection of $\mathcal{F}_i^\pi$ and $\mathcal{F}_j^\pi$ is isomorphic to $\mathcal{J}$. Note that then $t_{ij} = e(T_\mathcal{J})$ and $|\mathcal{K}_i \cup \mathcal{K}_j| = 2t - \kappa_\mathcal{J}$.

We will bound $M_\mathcal{J}$ by the number of (uncolored) copies of $\mathcal{F}_-^\pi \cup_{\mathcal{J}_-} \mathcal{F}_-^\pi$ which are created in the first round times the $\Theta(n^{2t-\kappa_\mathcal{J}})$ choices for $\mathcal{K}_i$ and $\mathcal{K}_j$ from all r-sets of the second round. Here $\mathcal{F}_-^\pi \cup_{\mathcal{J}_-} \mathcal{F}_-^\pi$ denotes an uncolored r-matched graph formed by the union of two copies of $\mathcal{F}_-^\pi$ which intersect in $\mathcal{J}_-$.

Let thus $M'_\mathcal{J}$ denote the number of copies of $\mathcal{F}_-^\pi \cup_{\mathcal{J}_-} \mathcal{F}_-^\pi$ contained in $G_{n/(2r)}$, multiplied with the number of choices for $\mathcal{K}_i$ and $\mathcal{K}_j$ from the r-sets of the second round. Note that $M'_\mathcal{J}$ is a random variable that depends only on the edges of the first round, and that $M_\mathcal{J} \leq M'_\mathcal{J}$. We have

$$\begin{aligned}
\mathbb{E}[M'_\mathcal{J}] &= \Theta(n^{2\kappa(\mathcal{F}_-^\pi)-\kappa(\mathcal{J}_-)} p^{2e(\mathcal{F}_-^\pi)-e(\mathcal{J}_-)}) \cdot \Theta(n^{2t-\kappa_\mathcal{J}}) \\
&= \Theta(n^{2\kappa(\mathcal{F}^\pi)} p^{2e(\mathcal{F}_-^\pi)}) n^{-\kappa(\mathcal{J}_-)-\kappa_\mathcal{J}} p^{-e(\mathcal{J}_-)} \\
&\overset{(6.31)}{=} \Theta(M^2) n^{-\kappa(\mathcal{J})} p^{-e(\mathcal{J}_-)} \\
&= \Theta(M^2) n^{-\kappa(\mathcal{J})} p^{-e(\mathcal{J})+e(T_\mathcal{J})}.
\end{aligned} \tag{6.35}$$

As $\mathcal{F}^\pi$ consists of t disjoint copies of F^π we can apply Lemma 6.19 once for each intersection of $\mathcal{J} \subseteq \mathcal{F}^\pi$ with one of the copies of F^π. For every such intersection $J \subseteq F^\pi$, as $p \gg n^{-\theta'}$, we have by Lemma 6.19 that

$$n^{-\kappa(J)} p^{-e(J)} \ll n^{-\kappa(J)+\theta e(J)} \overset{(6.6)}{=} n^{-\mu_{r,\theta}(J)} = \mathcal{O}(1).$$

As t is a fixed constant the same holds if we replace J by $\mathcal{J}$. Together with (6.35) and Markov's inequality this implies that a.a.s.

$$M'_\mathcal{J} \ll M^2 p^{e(T_\mathcal{J})} \tag{6.36}$$

(i.e., for an appropriate function $f(n) = o(1)$ a.a.s. it holds that $M'_\mathcal{J} \leq f(n) M^2 p^{e(T_\mathcal{J})}$). As the number of ways of choosing $\mathcal{J} \subseteq \mathcal{F}^\pi$ is a constant depending only on F, r and π, (6.36) holds a.a.s. for every possible choice of $\mathcal{J}$ simultaneously. For the second round we condition on the first one satisfying (6.30) and (6.36) for all $\mathcal{J} \subseteq \mathcal{F}^\pi$. With this we obtain from (6.34) that

$$\begin{aligned}\mathrm{Var}[Z] &= \sum_{(i,j)\in I} p^{2rt\deg_F(u_1)-t_{ij}} = \sum_{\substack{\mathcal{J}\subseteq\mathcal{F}^\pi:\\ \kappa_{\mathcal{J}}\geq 1}} M_{\mathcal{J}} p^{2rt\deg_F(u_1)-e(T_{\mathcal{J}})} \\ &\leq \sum_{\substack{\mathcal{J}\subseteq\mathcal{F}^\pi:\\ \kappa_{\mathcal{J}}\geq 1}} M'_{\mathcal{J}} p^{2rt\deg_F(u_1)-e(T_{\mathcal{J}})} \ll \left(Mp^{rt\deg_F(u_1)}\right)^2 \overset{(6.33)}{=} \mathbb{E}[Z]^2.\end{aligned}$$

By the second moment method this implies that a.a.s. $Z = \Theta\big(n^{\kappa(\mathcal{F}^\pi)}p^{e(\mathcal{F}^\pi)}\big)$, and that thus at least this number of copies of $\mathcal{F}^\pi$ are created in the second round. □

6.6.2 Proofs of the technical lemmas

The proofs in this section are essentially line-by-line translations of the analogous proofs in Mütze et al. [MST11] from the edge to the vertex case.

Together with Theorem 6.3 the following lemma proves Theorem 6.4.

Lemma 6.21. *Let F be a fixed nonempty graph, and let $r \geq 2$ be a fixed integer. Let $\theta^* = \theta^*(F,r)$ be the unique solution of*

$$\Lambda_{r,\theta}(F) \overset{!}{=} 0,$$

where $\Lambda_{r,\theta}(F)$ is defined in (6.4) *and* (6.5)*. Then we have*

$$m^{r*}(F) = \frac{1}{\theta^*(F,r)}.$$

Proof. For any nonempty ordered graph (F,π), set

$$\begin{aligned}\vec{\mathcal{H}}(F,\pi) := \big\{\vec{H} = \big((H_1,\sigma),H_2,\ldots,H_h\big) \mid H_1 \subseteq F \\ \wedge\, \sigma = \pi|_{H_1} = (u_1,\ldots,u_h) \\ \wedge\, \forall i\geq 2 : (H_i \subseteq H_1 \setminus \{u_1,\ldots,u_{i-1}\} \wedge u_i \in H_i)\big\}\end{aligned} \quad (6.37)$$

(cf. the maximizations in (6.2) and (6.3)). For all $\vec{H} \in \vec{\mathcal{H}}(F,\pi)$ we define

$$\begin{aligned} e^{r*}(\vec{H}) &:= \sum_{i=1}^{h} c_i e(H_i), \\ v^{r*}(\vec{H}) &:= 1 + \sum_{i=1}^{h} c_i \big(v(H_i) - 1\big), \end{aligned} \tag{6.38}$$

where the coefficients $c_i = c_i(\vec{H}, r)$ are defined as in (6.1). Furthermore, we define

$$\mu^*_{r,\theta}(\vec{H}) := v^{r*}(\vec{H}) - \theta \cdot e^{r*}(\vec{H}). \tag{6.39}$$

Note that by the definitions in (6.3) and (6.2), we have

$$\begin{aligned} m^{r*}(F,\pi) &= \min_{\pi \in \Pi(V(F))} \max_{H_1 \subseteq F} d^{r*}(H_1, \pi|_{H_1}) \\ &= \min_{\pi \in \Pi(V(F))} \max_{\vec{H} \in \vec{\mathcal{H}}(F,\pi)} \frac{e^{r*}(\vec{H})}{v^{r*}(\vec{H})} = \frac{1}{\theta^{**}(F,r)}, \end{aligned}$$

where $\theta^{**}(F,r)$ is the unique solution of

$$\max_{\pi \in \Pi(V(F))} \min_{\vec{H} \in \vec{\mathcal{H}}(F,\pi)} \mu^*_{r,\theta}(\vec{H}) \stackrel{!}{=} 0.$$

To prove Lemma 6.21, it suffices to show that the left hand side of the last equation equals $\Lambda_{r,\theta}(F)$ as defined in (6.5). We will do so by showing that for any nonempty ordered graph (F,π) and any $r \geq 2$ and $0 \leq \theta \leq 2$ we have

$$\min_{\vec{H} \in \vec{\mathcal{H}}(F,\pi)} \mu^*_{r,\theta}(\vec{H}) = \min_{H \subseteq F} \lambda_{r,\theta}(H, \pi|_H). \tag{6.40}$$

The remainder of the proof is devoted to establishing (6.40). To simplify the notation we consider r and θ fixed and drop all corresponding sub- and superscripts. In the following equations we define the quantities $\tilde{e}$, $\tilde{v}$ and $\tilde{\mu}$, which depend on the choice of an ordered graph (H_1, σ). In principle we should write $\tilde{e}_{(H_1,\sigma)}$, $\tilde{v}_{(H_1,\sigma)}$ and $\tilde{\mu}_{(H_1,\sigma)}$, but we omit this dependency from the notation as well.

Consider the following recursive definitions for $1 \le i \le h$:

$$\tilde{e}(H_i, \dots, H_h) := e(H_i) + (r-1) \cdot \sum_{j=i+1}^{h} \mathbf{1}_{\{u_j \in H_i\}} \tilde{e}(H_j, \dots, H_h)$$
$$\tilde{v}(H_i, \dots, H_h) := v(H_i) - 1 + (r-1) \cdot \sum_{j=i+1}^{h} \mathbf{1}_{\{u_j \in H_i\}} \tilde{v}(H_j, \dots, H_h). \tag{6.41}$$

We can now write $e^*(\vec{H})$ and $v^*(\vec{H})$ as

$$\begin{aligned} e^*(\vec{H}) &= r \cdot \tilde{e}(H_1, \dots, H_h) \\ v^*(\vec{H}) &= 1 + r \cdot \tilde{v}(H_1, \dots, H_h). \end{aligned} \tag{6.42}$$

This can be verified by induction, using the definition of c_i in (6.1) and noting that for $1 \le k \le h$ we have

$$e^*(\vec{H}) = \sum_{i=1}^{k} c_i e(H_i) + (r-1) \cdot \sum_{j=k+1}^{h} \Big(\sum_{i=1}^{k} c_i \mathbf{1}_{\{u_j \in H_i\}} \Big) \tilde{e}(H_j, \dots, H_h)$$

$$v^*(\vec{H}) = 1 + \sum_{i=1}^{k} c_i \big(v(H_i) - 1 \big)$$
$$\qquad + (r-1) \cdot \sum_{j=k+1}^{h} \Big(\sum_{i=1}^{k} c_i \mathbf{1}_{\{u_j \in H_i\}} \Big) \tilde{v}(H_j, \dots, H_h),$$

which is equivalent to (6.38) for $k = h$ and to (6.42) for $k = 1$. Combining (6.41) and (6.42) via (6.39) also yields that

$$\mu^*(\vec{H}) = 1 + r\tilde{\mu}(H_1, \dots, H_h), \tag{6.43}$$

where

$$\tilde{\mu}(H_i, \dots, H_h) := (v(H_i) - 1) - \theta e(H_i) + (r-1) \sum_{j=i+1}^{h} \mathbf{1}_{\{u_j \in H_i\}} \tilde{\mu}(H_j, \dots, H_h). \tag{6.44}$$

It follows that for any fixed subgraph $H_1 \subseteq F$ and $\sigma := \pi|_{H_1} = (u_1, \dots, u_h)$ the following holds: for $1 \le i \le h$ and any graph $H_i \subseteq H_1 \setminus \{u_1, \dots, u_{i-1}\}$ with $u_i \in H_i$, the value

$$\tilde{\lambda}_{(H_1,\sigma)}(H_i, i) := \min_{\substack{H_{i+1}, \dots, H_h \\ \forall j \ge i+1 : H_j \subseteq H_1 \setminus \{u_1, \dots, u_{j-1}\} \wedge u_j \in H_j}} \tilde{\mu}(H_i, \dots, H_h) \tag{6.45}$$

can be computed recursively via

$$\tilde{\lambda}_{(H_1,\sigma)}(H_i,i) = (v(H_i)-1) - \theta e(H_i) \\ + (r-1)\cdot \sum_{j=i+1}^{h} \mathbf{1}_{\{u_j\in H_i\}} \cdot \min_{H_j\subseteq H_1\setminus\{u_1,\dots,u_{j-1}\}:\, u_j\in H_j} \tilde{\lambda}_{(H_1,\sigma)}(H_j,j). \tag{6.46}$$

In the remainder of the proof we simplify the recursion on the right side to relate it to $\lambda()$ as defined in (6.4). First we show that we can get rid of the dependency on (H_1,σ), and that the value of $\tilde{\lambda}_{(H_1,\sigma)}(H_i,i)$ in fact only depends on the isomorphism class of $(H_i,\sigma|_{H_i})$. To this end, we prove that for any fixed ordered graph (H_1,σ) there exists a sequence $H_2,\dots,H_h\subseteq H_1$ as in (6.37) minimizing $\tilde{\mu}(H_1,\dots,H_h)$ with the additional property that

$$u_j\in H_i \Rightarrow H_j\subseteq H_i. \tag{6.47}$$

Let $H_2,\dots,H_h\subseteq H_1$ be graphs minimizing $\tilde{\mu}(H_1,\dots,H_h)$ such that every H_i is inclusion-maximal, and assume for the sake of contradiction that there exist indices $2\le i<j$ with $u_j\in H_i$ but $H_j\not\subseteq H_i$. Our choice of $H_2,\dots,H_h$ implies that for $H_i' := H_i\cup H_j$ and $H_j' := H_i\cap H_j$ we have

$$\tilde{\mu}(H_i',\dots,H_h) - \tilde{\mu}(H_i,\dots,H_h) > 0,$$
$$\tilde{\mu}(H_j,\dots,H_h) - \tilde{\mu}(H_j',\dots,H_h) \le 0,$$

where the first inequality is strict due to the inclusion-maximality of H_i. Expanding the above equations according to (6.44) yields that both terms are equal to

$$\big(v(H_j)-v(H_j')\big) - \theta\big(e(H_j)-e(H_j')\big) + \\ + (r-1)\sum_{k=j+1}^{h} \mathbf{1}_{\{u_k\in H_j\setminus H_i\}}\tilde{\mu}(H_k,\dots,H_h),$$

which is a contradiction. W.l.o.g. we may therefore assume that (6.47) holds, and that in (6.46) we can minimize over subgraphs of the graph $H_i\setminus\{u_i,\dots,u_{j-1}\}$ instead of subgraphs of $H_1\setminus\{u_1,\dots,u_{j-1}\}$.

Observe that in (6.46) the context (H_1,σ) is now irrelevant, and that we only require the ordering $\sigma|_{H_i}$ on the right hand side. Setting

$$\tilde{\lambda}_{(H_1,\sigma)}(H_i,i) =: \tilde{\lambda}(H_i,\sigma|_{H_i}). \tag{6.48}$$

and changing notations accordingly, we obtain from (6.46)

$$\tilde{\lambda}(H, \tau =: (u_1, \dots, u_h)) = \big(v(H) - 1\big) - \theta e(H) \\ + (r-1) \cdot \sum_{j=2}^{h} \min_{J \subseteq H \setminus \{u_1, \dots, u_{j-1}\} : u_j \in J} \tilde{\lambda}(J, \tau|_J). \quad (6.49)$$

Next we get rid of the sum in the equation above as follows:

$$\begin{aligned} \tilde{\lambda}(H, \tau) = {} & \big(v(H) - 1\big) - \theta e(H) \\ & + (r-1) \min_{J \subseteq H \setminus u_1 : u_2 \in J} \tilde{\lambda}(J, \tau|_J) \\ & + (r-1) \cdot \sum_{j=3}^{h} \min_{J \subseteq H \setminus \{u_1, \dots, u_{j-1}\} : u_j \in J} \tilde{\lambda}(J, \tau|_J). \\ = {} & (v(H \setminus u_1) - 1) + 1 - \theta e(H \setminus u_1) - \theta \deg_H(u_1) \\ & + (r-1) \min_{J \subseteq H \setminus u_1 : u_2 \in J} \tilde{\lambda}(J, \tau|_J) \\ & + (r-1) \cdot \sum_{j=3}^{h} \min_{J \subseteq H \setminus \{u_1, \dots, u_{j-1}\} : u_j \in J} \tilde{\lambda}(J, \tau|_J). \\ = {} & 1 + \tilde{\lambda}(H \setminus u_1, \tau|_{H \setminus u_1}) - \theta \deg_H(u_1) + (r-1) \min_{\substack{J \subseteq H \setminus u_1: \\ u_2 \in J}} \tilde{\lambda}(J, \tau|_J). \end{aligned} \quad (6.50)$$

Substituting

$$\begin{aligned} \tilde{\lambda}(H, \tau) &=: \bar{\lambda}(H \setminus u_1, \tau \setminus u_1) - \theta \deg_H(u_1), \\ (H \setminus u_1, \tau \setminus u_1) &=: (\bar{H}, \bar{\tau}), \\ u_2 &=: \bar{u}_1 \end{aligned} \quad (6.51)$$

we see that the last line of (6.50) is equivalent to

$$\bar{\lambda}(\bar{H}, \bar{\tau}) = 1 + \bar{\lambda}(\bar{H} \setminus \bar{u}_1, \bar{\tau} \setminus \bar{u}_1) - \theta \deg_{\bar{H}}(\bar{u}_1) \\ + (r-1) \min_{\substack{J \subseteq \bar{H}: \\ \bar{u}_1 \in J}} \Big(\bar{\lambda}(J \setminus \bar{u}_1, \bar{\tau}|_{J \setminus \bar{u}_1}) - \theta \deg_J(\bar{u}_1)\Big),$$

which is the recursive step in the definition of $\lambda()$ in (6.4). Moreover if $(\bar{H}, \bar{\tau}) = (H \setminus u_1, \tau \setminus u_1)$ contains no vertices (i.e. H is a graph on 1 vertex and therefore no edges) we have

$$\bar{\lambda}(\bar{H}, \bar{\tau}) \overset{(6.51)}{=} \tilde{\lambda}(H, \tau) + \theta \deg_H(u_1) = \tilde{\lambda}(H, \tau) \overset{(6.49)}{=} 0 = \lambda(\bar{H}, \bar{\tau}).$$

This takes care of the base case and implies that $\bar{\lambda}(H,\tau) = \lambda(H,\tau)$ for all ordered graphs (H,τ). Thus we have for every fixed (H_1,σ), $\sigma = (u_1,\ldots,u_h)$, that

$$\begin{aligned}\min_{\substack{H_2,\ldots,H_h\\ \forall j\geq 2: H_j\subseteq H_1\setminus\{u_1,\ldots,u_{j-1}\}\wedge u_j\in H_j}} \tilde{\mu}(H_1,\ldots,H_h) &\overset{(6.45)}{=} \tilde{\lambda}_{(H_1,\sigma)}(H_1,1)\\ &\overset{(6.48)}{=} \tilde{\lambda}(H_1,\sigma) \overset{(6.51)}{=} \bar{\lambda}(H_1\setminus u_1,\sigma\setminus u_1) - \theta \deg_{H_1}(u_1)\\ &= \lambda(H_1\setminus u_1,\sigma\setminus u_1) - \theta\deg_{H_1}(u_1). \qquad (6.52)\end{aligned}$$

Still using the notation $\pi|_{H_1} = \sigma = (u_1,\ldots,u_h)$ (cf. (6.37)), equation (6.40) now follows from

$$\begin{aligned}&\min_{\vec{H}\in\vec{\mathcal{H}}(F,\pi)} \mu^*_{r,\theta}(\vec{H})\\ &\overset{(6.37),(6.43)}{=} \min_{H_1\subseteq F}\Big\{1 + r\cdot \min_{\substack{H_2,\ldots,H_h\\ \forall j\geq 2: H_j\subseteq H_1\setminus\{u_1,\ldots,u_{j-1}\}\wedge u_j\in H_j}} \tilde{\mu}(H_1,\ldots,H_h)\Big\}\\ &\overset{(6.52)}{=} \min_{H_1\subseteq F}\Big\{1 + r\big(\lambda(H_1\setminus u_1,\sigma\setminus u_1) - \theta\deg_{H_1}(u_1)\big)\Big\}\\ &\overset{(6.12)}{=} \min_{H_1\subseteq F}\lambda_{r,\theta}(H_1,\sigma) = \min_{H\subseteq F}\lambda_{r,\theta}(H,\pi|_H),\end{aligned}$$

where in the last step we applied Lemma 6.9 to the family of all ordered subgraphs of (F,π). □

It remains to prove Lemma 6.19.

Proof of Lemma 6.19. For this proof we require an extension of the definition of connectedness to r-matched graphs. We call an r-matched graph $H = (V,E,\mathcal{K})$ *connected* if for any 2 vertices $u,v\in V$ which are not part of the same r-set, there exists a sequence of r-sets $K_1,\ldots,K_t\in\mathcal{K}$ such that $u\in K_1$, $v\in K_t$ and there exists an edge between at least one vertex in K_i and one in K_{i+1} for all $1\leq i\leq t-1$. Since the value of $\mu_{r,\theta}()$ for a disconnected r-matched graph H is simply the sum of the values of $\mu_{r,\theta}()$ for all connected components of H, it suffices to prove the claim for all connected r-matched subgraphs $J\subseteq F_r^\pi$, i.e. to prove that for any integer $r\geq 2$ and any $0\leq\theta\leq 2$ we have

$$\min_{\substack{J\subseteq F_r^\pi:\\ J\text{ connected}}} \mu_{r,\theta}(J) = \min_{H\subseteq F}\lambda_{r,\theta}(H,\pi|_H). \qquad (6.53)$$

For the remainder of the proof we consider r and θ fixed and drop the corresponding subscripts from the notation.

Let $\pi = (u_1, \dots, u_f)$, and define $F_{i-} := F \setminus \{u_1, \dots, u_i\}$ and $\pi_{i-} := \pi|_{F_{i-}}$. For any grey-black r-matched graph $(F_{i-})^{\pi_{i-}}$ we call the r-set containing the vertex u_{i+1} of its central copy of F_{i-} the *central r-set*.

Let J be a connected subgraph of F^π, and let $0 \leq i \leq f-1$ be the largest index such that J is also contained in a copy of $(F_{i-})^{\pi_{i-}}$. By the maximal choice of i and the connectedness of J, the graph J contains the central r-set of this copy. With this we can reformulate equation (6.53) to

$$\min_{0 \leq i \leq f-1} \min_{\substack{J \subseteq (F_{i-})^{\pi_{i-}}: \\ K(u_{i+1}) \in J \wedge J \text{ connected}}} \mu(J) = \min_{0 \leq i \leq f-1} \min_{\substack{H \subseteq F_{i-}: \\ u_{i+1} \in H}} \lambda(H, \pi|_H). \tag{6.54}$$

where we use $K(u_{i+1}) \in J$ as a shorthand notation to indicate that J contains the central r-set $K(u_{i+1})$ of $(F_{i-})^{\pi_{i-}}$. We now show that the inner minimizations of (6.54) are equivalent. By changing variables ($F \leftarrow F_{i-}$ and $\pi \leftarrow \pi_i$) this reduces to showing that for any ordered graph (F, π) we have

$$\min_{\substack{J \subseteq F^\pi: K(u_1) \in J \\ J \text{ connected}}} \mu(J) = \min_{\substack{H \subseteq F: \\ u_1 \in H}} \lambda(H, \pi|_H). \tag{6.55}$$

For any r-matched graph H we refer to a subgraph $J \subseteq H$ that minimizes $\mu(J)$ as a *rarest subgraph* of H. To determine a rarest subgraph of F^π we can make use of its recursive structure.

Let $1 \leq i \leq f-1$, and consider a fixed copy $(\hat{F}_{i-})^{\pi_{i-}}$ of $(F_{i-})^{\pi_{i-}}$ in F^π. By $\hat{F}_{i-}$ we denote the central copy of (F_{i-}, π_{i-}) in $(\hat{F}_{i-})^{\pi_{i-}}$, and by $\hat{u}_i$ the vertex that completes $\hat{F}_{i-}$ to a copy of $(F_{(i-1)-}, \pi_{(i-1)-})$. Moreover, let $(\hat{F}_{(i-1)-})^{\pi_{(i-1)-}}$ denote the copy of $(F_{(i-1)-})^{\pi_{(i-1)-}}$ that is formed by $(\hat{F}_{i-})^{\pi_{i-}}$, $K(\hat{u}_i)$ and $r-1$ other copies of $(F_{i-})^{\pi_{i-}}$.

Note that the r copies of $(F_{i-})^{\pi_{i-}}$ joined at the central r-set $K(\hat{u}_i) = (u_{i,1}, \dots, u_{i,r})$ of $(\hat{F}_{(i-1)-})^{\pi_{(i-1)-}}$ are essentially independent: For each vertex $u_{i,\ell}$, $1 \leq \ell \leq r$, we consider the graph obtained by removing from $(\hat{F}_{(i-1)-})^{\pi_{(i-1)-}}$ the $r-1$ copies of $(F_{i-})^{\pi_{i-}}$ that are *not* associated with $u_{i,\ell}$, i.e., whose central copy of (F_{i-}, π_{i-}) is *not* connected by $\deg_{F_{(i-1)-}}(u_i)$ edges to $u_{i,\ell}$. (If $\deg_{F_{(i-1)-}}(u_i) = 0$ we associate the copies of $(F_{i-})^{\pi_{i-}}$ with the vertices of $K(\hat{u}_i)$ arbitrarily.) We call this graph the *branch* of $(\hat{F}_{(i-1)-})^{\pi_{(i-1)-}}$ corresponding to $u_{i,\ell}$. Note that this is still an r-matched graph and that all r branches contain the central r-set $K(\hat{u}_i)$. By the linearity of $\mu(H)$ in $e(H)$ and $\kappa(H)$, a rarest connected subgraph of $(\hat{F}_{(i-1)-})^{\pi(i-1)-}$ containing $K(\hat{u}_i)$ can be found by determining a rarest connected subgraph containing $K(\hat{u}_i)$ in each branch of $(\hat{F}_{(i-1)-})^{\pi(i-1)-}$ independently. Let $\hat{J}_i$ denote an arbitrary fixed such rarest subgraph. Note that in particular we can compute $\mu(J)$ as on the left hand side of (6.55) as

$$\mu(J) = 1 + r(\mu(\hat{J}_1) - 1). \tag{6.56}$$

Similarly, we can find a rarest connected subgraph J_i containing $K(\hat{u}_i)$ for a branch of $(\hat{F}_{(i-1)-})^{\pi(i-1)-}$ by determining an optimal choice for $H_i := J_i \cap \hat{F}_{(i-1)-}$. For any choice of H_i, by recursion, for each vertex u_j' of H_i, $i+1 \leq j \leq f$, we already know a rarest subgraph containing $K(u_j')$ for all the $r-1$ remaining branches of the copy of $(F_{(j-1)-})^{\pi(j-1)-}$ corresponding to the other $r-1$ vertices of $K(u_j')$. Letting $\hat{J}_j$, $i+1 \leq j \leq f$ denote such rarest subgraphs, we obtain that the value of $\mu(J_i)$ resulting from a given choice of $H_i \subseteq \hat{F}_{(i-1)-}$ is

$$\mu(J_i) = v(H_i) - \theta e(H_i) + \sum_{j=i+1}^{f} \mathbf{1}_{\{u_i \in H_j\}} (r-1)(\mu(\hat{J}_j) - 1).$$

Here we used that for $i+1 \leq j \leq f$ all $r-1$ many copies of $\hat{J}_j$ share one r-set, and that each such r-set also contains one vertex of $v(H_i)$.

Substituting $\mu(J_i) - 1 =: \tilde{\lambda}_{(F,\pi)}(H_i, i)$ in the above equation yields for $1 \leq i \leq f$ the recursion

$$\begin{aligned}\tilde{\lambda}_{(F,\pi)}(H_i, i) &= (v(H_i) - 1) - \theta e(H_i) \\ &\qquad + (r-1) \sum_{j=i+1}^{f} \mathbf{1}_{\{u_j \in H_i\}} \min_{\substack{H_j \subseteq H_1 \setminus \{u_1, \ldots, u_{j-1}\}: \\ u_j \in H_j}} \tilde{\lambda}_{(F,\pi)}(H_j, j).\end{aligned}$$

This is essentially the same recursion as (6.46) in the proof of Lemma 6.21. Analogously to the proof of Lemma 6.21 one can show that

$$\tilde{\lambda}_{(F,\pi)}(H_1, 1) = \lambda(H_1 \setminus u_1, \sigma \setminus u_1) - \theta \deg_{H_1}(u_1), \tag{6.57}$$

where $\sigma := \pi|_{H_1}$ (cf. (6.52)). Finally

$$\begin{aligned}\min_{\substack{J \subseteq F^\pi : K(u_1) \in J \\ \wedge J \text{ connected}}} \mu(J) &\overset{(6.56)}{=} r\big(\mu(\hat{J}_1) - 1\big) + 1 = \min_{H_1 \subseteq F : u_1 \in H} 1 + r \cdot \tilde{\lambda}(H_1, 1) \\ &\overset{(6.57)}{=} \min_{H_1 \subseteq F : u_1 \in H_1} 1 + r \cdot \big(\lambda(H_1 \setminus u_1, \sigma \setminus u_1) - \theta \deg_{H_1}(u_1)\big) \\ &\overset{(6.12)}{=} \min_{H_1 \subseteq F : u_1 \in H_1} \lambda(H_1, \sigma) = \min_{H \subseteq F : u_1 \in H} \lambda(H, \pi|_H).\end{aligned}$$

In the last line we applied Lemma 6.9 to the family of all ordered subgraphs of (F, π) that contain the vertex u_1. This shows (6.55) and finishes the proof. $\square$

Bibliography

[ABHP13] Peter Allen, Julia Böttcher, Jan Hladký, and Diana Piguet, *Turánnical hypergraphs*, Random Structures & Algorithms **42** (2013), no. 1, 29–58.

[ADS09] Dimitris Achlioptas, Raissa M. D'Souza, and Joel H. Spencer, *Explosive percolation in random networks*, Science **323** (2009), no. 5920, 1453–1455.

[BB10] József Balogh and Jane Butterfield, *Online Ramsey games for triangles in random graphs*, Discrete Mathematics **310** (2010), no. 24, 3653–3657.

[BF01] Tom Bohman and Alan Frieze, *Avoiding a giant component*, Random Structures & Algorithms **19** (2001), no. 1, 75–85.

[BMS12a] József Balogh, Robert Morris, and Wojciech Samotij, *Independent sets in hypergraphs*, arXiv preprint arXiv:1204.6530 (2012).

[BMS12b] Michael Belfrage, Torsten Mütze, and Reto Spöhel, *Probabilistic one-player Ramsey games via deterministic two-player games*, SIAM Journal on Discrete Mathematics **26** (2012), no. 3, 1031–1049.

[Bol81] Béla Bollobás, *Threshold functions for small subgraphs*, Mathematical Proceedings of the Cambridge Philosophical Society **90** (1981), no. 2, 197–206.

[BT87] Béla Bollobás and Andrew Thomason, *Threshold functions*, Combinatorica **7** (1987), no. 1, 35–38.

[CG10] David Conlon and William Timothy Gowers, *Combinatorial theorems in sparse random sets*, arXiv preprint arXiv:1011.4310 (2010).

[Con09] David Conlon, *A new upper bound for diagonal Ramsey numbers*, Annals of Mathematics **170** (2009), no. 2, 941–960.

[EFRS78] Paul Erdős, Ralph J. Faudree, Cecil C. Rousseau, and Richard H. Schelp, *The size Ramsey number*, Periodica Mathematica Hungarica **9** (1978), no. 1–2, 145–161.

[ER59] Paul Erdős and Alfréd Rényi, *On random graphs, I*, Publicationes Mathematicae **6** (1959), 290–297.

[ER60] ———, *On the evolution of random graphs*, Magyar Tudományos Akadémia Matematikai Kutató Intézet Közleményei **5** (1960), 17–61.

[ER61a] ———, *On the evolution of random graphs*, Bulletin de L'Institut International de Statistique **38** (1961), no. 4, 343–347.

[ER61b] ———, *On the strength of connectedness of a random graph*, Acta Mathematica Hungarica **12** (1961), no. 1, 261–267.

[ER63] ———, *Asymmetric graphs*, Acta Mathematica Hungarica **14** (1963), no. 3, 295–315.

[ER64] ———, *On random matrices*, Magyar Tudományos Akadémia Matematikai Kutató Intézet Közleményei **8** (1964), 455–461.

[ER66] ———, *On the existence of a factor of degree one of a connected random graph*, Acta Mathematica Hungarica **17** (1966), no. 3, 359–368.

[ER68] ———, *On random matrices II*, Studia Scientiarum Mathematicarum Hungarica **3** (1968), 459–464.

[Erd47] Paul Erdős, *Some remarks on the theory of graphs*, Bulletin of the American Mathematical Society **53** (1947), no. 4, 292–294.

[ES35] Paul Erdős and George Szekeres, *A combinatorial problem in geometry*, Compositio Mathematica **2** (1935), 463–470.

[Exo89] Geoffrey Exoo, *A lower bound for $R(5,5)$*, Journal of graph theory **13** (1989), no. 1, 97–98.

[FK00] Ehud Friedgut and Michael Krivelevich, *Sharp thresholds for certain Ramsey properties of random graphs*, Random Structures & Algorithms **17** (2000), no. 1, 1–19.

[FKR+03] Ehud Friedgut, Yoshiharu Kohayakawa, Vojtěch Rödl, Andrzej Ruciński, and Prasad Tetali, *Ramsey games against a one-armed bandit*, Combinatorics, Probability and Computing **12** (2003), no. 5–6, 515–545.

[FRRT06] Ehud Friedgut, Vojtěch Rödl, Andrzej Ruciński, and Prasad Tetali, *A sharp threshold for random graphs with a monochromatic triangle in every edge coloring*, Memoirs of the American Mathematical Society **179** (2006), no. 845, 66 pp.

[FRS10] Ehud Friedgut, Vojtěch Rödl, and Mathias Schacht, *Ramsey properties of random discrete structures*, Random Structures & Algorithms **37** (2010), no. 4, 407–436.

[Gil59] Edgar N Gilbert, *Random graphs*, The Annals of Mathematical Statistics **30** (1959), no. 4, 1141–1144.

[GPST11] Luca Gugelmann, Yury Person, Angelika Steger, and Henning Thomas, *A randomized version of Ramsey's theorem*, Electronic Notes in Discrete Mathematics **38** (2011), 431–436.

[GPST12] ———, *A randomized version of Ramsey's theorem*, Random Structures & Algorithms **41** (2012), no. 4, 488–505.

[GR87] Ronald L. Graham and Vojtěch Rödl, *Numbers in Ramsey theory*, Surveys in Combinatorics 1987, London Mathematical Society Lecture Note Series, no. 123, Cambridge University Press, 1987, pp. 111–153.

[GRS90] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer, *Ramsey theory*, second ed., Wiley-Interscience, 1990.

[GS90] Ronald L. Graham and Joel H. Spencer, *Ramsey theory*, Scientific American **262** (1990), no. 7, 112–117.

[GS11] Luca Gugelmann and Reto Spöhel, *On balanced coloring games in random graphs*, Electronic Notes in Discrete Mathematics **38** (2011), 425–430.

[GS14] Luca Gugelmann and Reto Spöhel, *On balanced coloring games in random graphs*, European Journal of Combinatorics **35** (2014), no. 0, 297–312, Selected Papers of EuroComb'11.

[HJ63] Alfred W. Hales and Robert I. Jewett, *Regularity and positional games*, Transactions of the American Mathematical Society **106** (1963), 222–229.

[JŁR00] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński, *Random graphs*, Wiley-Interscience, New York, 2000.

[KK97] Yoshiharu Kohayakawa and Bernd Kreuter, *Threshold functions for asymmetric Ramsey properties involving cycles*, Random Structures & Algorithms **11** (1997), no. 3, 245–276.

[KŁR97] Yoshiharu Kohayakawa, Tomasz Łuczak, and Vojtech Roedl, *On K^4-free subgraphs of random graphs*, Combinatorica **17** (1997), no. 2, 173–213.

[KLS09] Michael Krivelevich, Po-Shen Loh, and Benny Sudakov, *Avoiding small subgraphs in Achlioptas processes*, Random Structures & Algorithms **34** (2009), no. 1, 165–195.

[KR83] Michał Karoński and Andrzej Ruciński, *On the number of strictly balanced subgraphs of a random graph*, Graph Theory (M. Borowiecki, John W. Kennedy, and Maciej M. Sysło, eds.), Lecture Notes in Mathematics, vol. 1018, Springer Berlin Heidelberg, 1983, pp. 79–83.

[KR97] Michał Karoński and Andrzej Ruciński, *The origins of the theory of random graphs*, The mathematics of Paul Erdős, I, Algorithms and Combinatorics, vol. 13, Springer, Berlin, 1997, pp. 311–336.

[Kre96] Bernd Kreuter, *Threshold functions for asymmetric Ramsey properties with respect to vertex colorings*, Random Structures & Algorithms **9** (1996), no. 3, 335–348.

[KSS] Yoshiharu Kohayakawa, Mathias Schacht, and Reto Spöhel, *Upper bounds on probability thresholds for asymmetric Ramsey properties*, Random Structures & Algorithms, to appear.

[KSS10] Michael Krivelevich, Reto Spöhel, and Angelika Steger, *Offline thresholds for Ramsey-type games on random graphs*, Random Structures & Algorithms **36** (2010), no. 1, 57–79.

[ŁRV92] Tomasz Łuczak, Andrzej Ruciński, and Bernd Voigt, *Ramsey properties of random graphs*, Journal of Combinatorial Theory, Series B **56** (1992), no. 1, 55–68.

[MMS07] Martin Marciniszyn, Dieter Mitsche, and Miloš Stojaković, *Online balanced graph avoidance games*, European Journal of Combinatorics **28** (2007), no. 8, 2248–2263.

[MR97] Brendan D McKay and Stanisław P Radziszowski, *Subgraph counting identities and Ramsey numbers*, Journal of Combinatorial Theory, Series B **69** (1997), no. 2, 193–209.

[MRS11] Torsten Mütze, Thomas Rast, and Reto Spöhel, *Coloring random graphs online without creating monochromatic subgraphs*, Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'11), SIAM, 2011, pp. 145–158.

[MS07] Martin Marciniszyn and Reto Spöhel, *Online vertex colorings of random graphs without monochromatic subgraphs*, Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2007, pp. 486–493.

[MSS05] Martin Marciniszyn, Reto Spöhel, and Angelika Steger, *The online clique avoidance game on random graphs*, Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques (Chandra Chekuri, Klaus Jansen, José D. P. Rolim, and Luca Trevisan, eds.), Lecture Notes in Computer Science, vol. 3624, Springer Berlin Heidelberg, 2005, pp. 390–401.

[MSS09a] Martin Marciniszyn, Reto Spöhel, and Angelika Steger, *Online Ramsey games in random graphs*, Combinatorics, Probability and Computing **18** (2009), no. 1–2, 271–300.

[MSS09b] ______, *Upper bounds for online Ramsey games in random graphs*, Combinatorics, Probability and Computing **18** (2009), no. 1–2, 259–270.

[MSSS09] Martin Marciniszyn, Jozef Skokan, Reto Spöhel, and Angelika Steger, *Asymmetric Ramsey properties of random graphs involving cliques*, Random Structures & Algorithms **34** (2009), no. 4, 419–453.

[MST11] Torsten Mütze, Reto Spöhel, and Henning Thomas, *Small subgraphs in random graphs and the power of multiple choices*, Journal of Combinatorial Theory, Series B **101** (2011), no. 4, 237–268.

[Nen13] Rajko Nenadov, personal communication, 2013.

[Noe12] Andreas Noever, *Online Ramsey games for triangles in random graphs*, Master's thesis, ETH Zurich, Switzerland, 2012.

[PST09] Anupam Prakash, Reto Spöhel, and Henning Thomas, *Balanced online Ramsey games in random graphs*, The Electronic Journal of Combinatorics **16** (2009), no. 1, R11.

[Rad33] Richard Rado, *Studien zur Kombinatorik*, Mathematische Zeitschrift **36** (1933), no. 1, 424–470.

[Rad11] Stanisław P. Radziszowski, *Small Ramsey numbers, dynamic survey*, The Electronic Journal of Combinatorics (2011).

[Ram30] Frank P. Ramsey, *On a problem of formal logic*, Proceedings of the London Mathematical Society **s2–30** (1930), no. 1, 264–286.

[RR93] Vojtěch Rödl and Andrzej Ruciński, *Lower bounds on probability thresholds for Ramsey properties*, Combinatorics, Paul Erdős is eighty, 1, Bolyai Society Mathematical Studies, vol. 1, János Bolyai Mathematical Society, 1993, pp. 317–346.

[RR94] ______, *Random graphs with monochromatic triangles in every edge coloring*, Random Structures & Algorithms **5** (1994), no. 2, 253–270.

[RR95] ______, *Threshold functions for Ramsey properties*, Journal of the American Mathematical Society **8** (1995), no. 4, 917–942.

[RR98] ______, *Ramsey properties of random hypergraphs*, Journal of Combinatorial Theory, Series A **81** (1998), no. 1, 1–33.

[RRS07] Vojtěch Rödl, Andrzej Ruciński, and Mathias Schacht, *Ramsey properties of random k-partite, k-uniform hypergraphs*, SIAM Journal on Discrete Mathematics **21** (2007), no. 2, 442–460.

[RW12] Oliver Riordan and Lutz Warnke, *Achlioptas process phase transitions are continuous*, The Annals of Applied Probability **22** (2012), no. 4, 1450–1464.

[Sch16] Issai Schur, *Über die Kongruenz* $x^m + y^m \equiv z^m (\bmod p)$, Jahresbericht der Deutschen Mathematiker-Vereinigung **25** (1916), 114–117.

[Spe75] Joel Spencer, *Ramsey's theorem – a new lower bound*, Journal of Combinatorial Theory. Series A **18** (1975), no. 1, 108–115.

[ST12] David Saxton and Andrew Thomason, *Hypergraph containers*, arXiv preprint arXiv:1204.6595 (2012).

[Sze75] E Szemerédi, *On sets of integers containing k elements in arithmetic progression*, Acta Arithmetica **27** (1975), no. 1, 199–245.

[Tho88] Andrew Thomason, *An upper bound for some Ramsey numbers*, Journal of Graph Theory **12** (1988), no. 4, 509–517.

[Tho13] Henning Thomas, *Aspects of games on random graphs*, Ph.D. thesis, ETH Zurich, 2013.

[Tur41] Pál Turán, *Eine extremalaufgabe aus der graphentheorie*, Matematikai és Fizikai Lapok **48** (1941), 436–452.

[vdW27] Bartel L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Archief voor Wiskunde **15** (1927), 212–216.